9-22-2014 12:00 AM

Visual Perception and Cognition in Image-Guided Intervention

Kamyar Abhari
The University of Western Ontario

Supervisor
Dr. Terry Peters
The University of Western Ontario Joint Supervisor
Dr. Roy Eagleson
The University of Western Ontario

Graduate Program in Biomedical Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

VISUAL PERCEPTION AND COGNITION IN IMAGE-GUIDED INTERVENTION

(Thesis format: Integrated Article)

by

Kamyar Abhari

Graduate Program in Biomedical Engineering

A thesis submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

The School of Graduate and Postdoctoral Studies

Western University

London, Ontario, Canada

Abstract

Surgical image visualization and interaction systems can dramatically affect the efficacy and efficiency of surgical training, planning, and interventions. This is even more profound in the case of minimally-invasive surgery where restricted access to the operative field in conjunction with limited field of view necessitate a visualization medium to provide patient-specific information at any given moment. Unfortunately, little research has been devoted to studying human factors associated with medical image displays and the need for a robust, intuitive visualization and interaction interfaces has remained largely unfulfilled to this day. Failure to engineer efficient medical solutions and design intuitive visualization interfaces is argued to be one of the major barriers to the meaningful transfer of innovative technology to the operating room. This thesis was, therefore, motivated by the need to study various cognitive and perceptual aspects of human factors in surgical image visualization systems, to increase the efficiency and effectiveness of medical interfaces, and ultimately to improve patient outcomes. To this end, we chose four different minimally-invasive interventions in the realm of surgical training, planning, training for planning, and navigation: The first chapter involves the use of stereoendoscopes to reduce morbidity in endoscopic third ventriculostomy. The results of this study suggest that, compared with conventional endoscopes, the detection of the basilar artery on the surface of the third ventricle can be facilitated with the use of stereoendoscopes, increasing the safety of targeting in third ventriculostomy procedures. In the second chapter, a contour enhancement technique is described to improve preoperative planning of arteriovenous malformation interventions. The proposed method, particularly when combined with stereopsis, is shown to increase the speed and accuracy of understanding the spatial relationship between vascular structures. In the third chapter, an augmented-reality system is proposed to facilitate the training of planning brain tumour resection. The results of our user study indicate that the proposed system improves subjects' performance, particularly novices', in formulat-

ing the optimal point of entry and surgical path independent of the sensorimotor tasks performed. In the last chapter, the role of fully-immersive simulation environments on the surgeons' non-technical skills to perform vertebroplasty procedure is investigated. Our results suggest that while training surgeons may increase their technical skills, the introduction of crisis scenarios significantly disturbs the performance, emphasizing the need of realistic simulation environments as part of training curriculum.

Acknowlegements

I would like to start by acknowledging my supervisor and my mentor, Dr. Terry Peters, whose expertise, understanding, and patience, added considerably to my graduate experience. Throughout my PhD, he has always been a source of guidance and support, and I truly feel honoured to have had the opportunity to learn from him and to benefit from his breadth of knowledge and deep insights.

I would like to express my sincere gratitude to my co-supervisor, Dr. Roy Eagleson, as well as Dr. Sandrine de Ribaupierre for their support in and out of graduate school. I am truly thankful for their help in writing papers and conducting experiments, and for their assistance at all levels of this research project. This work would have not been possible without them.

Special thanks goes out to John Moore, Chris Wedlake, Dr. Elvis Chen, Dr. Ali Khan for sharing their expertise and providing advice and assistance at every level of this research work. It has truly been a privilege to work with and learn from them. I would also like to thank Dr. David Holdsworth for his valuable advice and guidance along the way.

I am particularly grateful to John Baxter for his assistant, technical support, and his significant contribution to my studies. I would also like to thank Saeed Bakhshmand for his assistant in data collection.

I would like to thank Dr. Nassir Navab for the opportunity of working at his wonderful lab in Munich. I like to thank everybody at NARVIS lab, particularly Dr. Pascal Fallavolitta, Patrick Wucherer, Philipp Stefan, Dr. Matthias Weigl, and Dr. Simon Weidert for making my stay memorable.

I would also like to thank my examiners, Dr. Hanif Ladak, Dr. Tim Wilson, Dr. Ana Luisa Trejos, and Dr. Eddie Edwards for their time and for their helpful suggestions.

I would like to thank everybody in Dr. Peter's, Dr. Fenster's, and Dr. Drangova's lab, as well as many others at Robarts for their thoughtful discussions, helpful criticisms, and, most of all, for their friendship. I also like to thank all the surgeons, residents, and students who participated in my studies which sometimes lasted for more than three hours!

Finally, I would like to thank my parents and my brothers for their unconditional love and support. I am truly grateful for having them by me and none of my accomplishments would have been possible without them.

Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations

2AFC: 2-Alternative Forced Choice

AR: Augmented-Reality

AVM: Arteriovenous Malformations

AV: Augmented-Virtuality

CL: Cognitive Load

CSF: Cerebrospinal Fluid

CTA: Computed Tomography Angiography

CT: Computed Tomography

DVR: Direct Volume Rendering

DTI: Diffusion Tensor Imaging

ETV: Endoscopic Third Ventriculostomy

fMRI: function Magnetic Resonance Imaging

HCI: Human-Computer Interaction

HMD: Head-Mounted Display

IGI: Image-Guided Intervention

JND: Just Noticeable Difference

MIP: Maximum-Intensity Projection

MIS: Minimally-Invasive Surgery

MPR: Multi-Planner Reformatted

MR: Mixed-Reality

MRA: Magnetic Resonance Angiography

MRI: Magnetic Resonance Imaging

NPR: Non-Photo-Realistic

OR: Operation Room

PMMA: Polymethylmethacrylate

RT: Response Time

TF: Transfer Function

VR: Virtual-Reality

VS: Visuospatial Sketchpad

VE: Virtual Environments

XP: Crossed Plane

Chapter 1

Introduction

"The world, as we perceive it, is our own invention." *- Heinz von Foerster*

In minimally-invasive interventions, the target anatomy is often accessed by means of tools placed through small incisions in the body. Patients who undergo minimally-invasive interventions often have fewer complications, shorter hospital stays, quicker recovery times, and less pain than traditional open procedures. However, these benefits come at the cost of limited view of the surgical site and restricted access inside the body cavity. To overcome these issues, over the past several decades significant efforts have been dedicated to the developments in medical imaging, tracking and localization technologies, surgical navigation and simulation, image segmentation, modeling, and registration [1]. In the meantime, methods of visualization and interaction have undergone dramatic changes to permit rapid and accurate examination of medical images with minimal mental and physical effort. These advancements are made possible partly by improvements in the display technology, and partly by studying the role of human perception, cognition, and action in perceiving, processing, and interacting with medical images. In fact, researchers collectively agree that even the most technically optimal visualization and interaction environments would not optimally convey information if the information they deliver surpasses the users' perceptual and cognitive capabilities

[2].

In the realm of minimally-invasive interventions, *visual perception* is the capacity to transform preoperative and intraoperative images into a form that can be used by the *cognitive* system, along with domain knowledge, to formulate plans of *action*, accomplishing surgical tasks. This continual cycle of visual perception-cognition-action influences every aspect of minimally-invasive interventions, from training and preoperative planning to intraoperative navigation. This introductory chapter reviews some important aspects of the human visual system and its underlying mechanisms.

1.1 Visual Perception

Perception is the process by which we organize, recognize, and interpret sensory information to understand our surrounding environment. It begins when a *distal stimulus* emits or reflects energy stimulating a sensory organ. The energy falling on the receptors of the sensory organ, commonly referred to as *proximal stimulus*, generates neural activity which subsequently is transmitted to the brain. The process of creating a mental representation of the distal stimulus using the information contained in the proximal stimulus is described as *perception*. If the sensory neurons are activated by means of light, the process of constructing a mental image of the distal object is described as *visual perception* - or what we think of as *seeing*. This section is an introduction to visual perception and its underlying physiological and behavioural mechanisms.

1.1.1 Visual System

Perceiving a visual scene begins with reception of light by *photoreceptor* cells located in the layer of the retina, commonly known as *rods* and *cones* (Figure 1.1). The photoreceptor cells are connected to *ganglion* cells in which a biological conversion of the photons generates a burst of action potentials. The resultant signal is then transferred

to the *lateral geniculate nucleus* (LGN), a six-layered structure in the thalamus (Figure 1.2). The LGN consists of two major types of cells: *Parvo* and *Magno*. These cells behave differently in response to four major properties in an image, namely, *colour*, *motion*, *spatial resolution*, and *contrast sensitivity* [3]. Magno cells are effectively colour-blind, have a high contrast-sensitivity, and exhibit a fast response to temporal aspects of a visual stimulus. Furthermore, their receptive fields are 2 to 3 times larger than Parvo cells, making them less sensitive to high spatial frequencies. Parvo cells behave in exactly the opposite manner. These characteristics of Magno and Parvo cells are perpetuated in the primary visual cortex in the occipital lobe - or V1, contributing to different aspects of vision. Neurons contained in V1 are tuned to simple properties of the visual scene including *colour*, *form*, and *movement*. V1 projects mainly to V2 in which more complex properties such as *horizontal disparity* is detected [3].

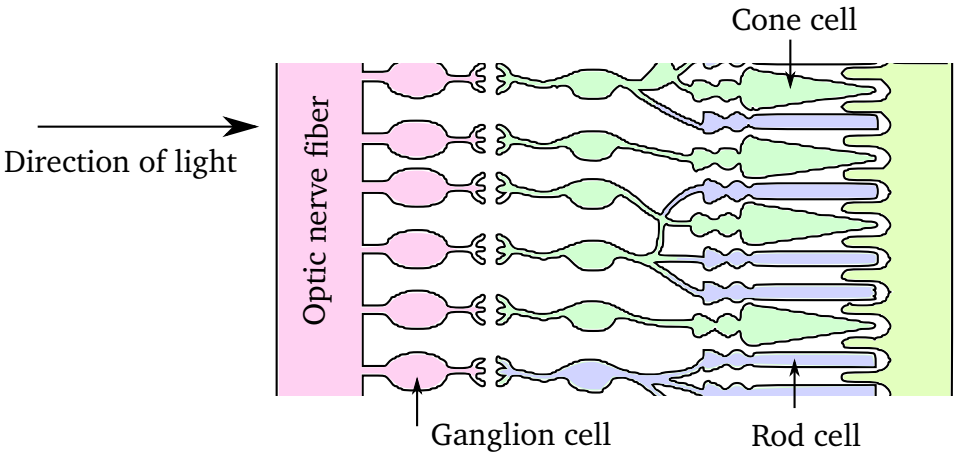


Figure 1.1: Cross-section of main retinal layers

As visual information exits V1 and V2, it travels via two separate pathways: the *dorsal stream*, commonly known as *where* pathway or *how* pathway, that runs forward from V1 into the posterior parietal lobe through V3, MT (also known as V5), and MST brain areas; and the *ventral stream*, or *what* pathway, that stretches from V1 downward into the inferior temporal lobe through V4 and IT visual areas (Figure 1.3). The dorsal stream is associated with motion, depth, and spatial organization. Colour perception

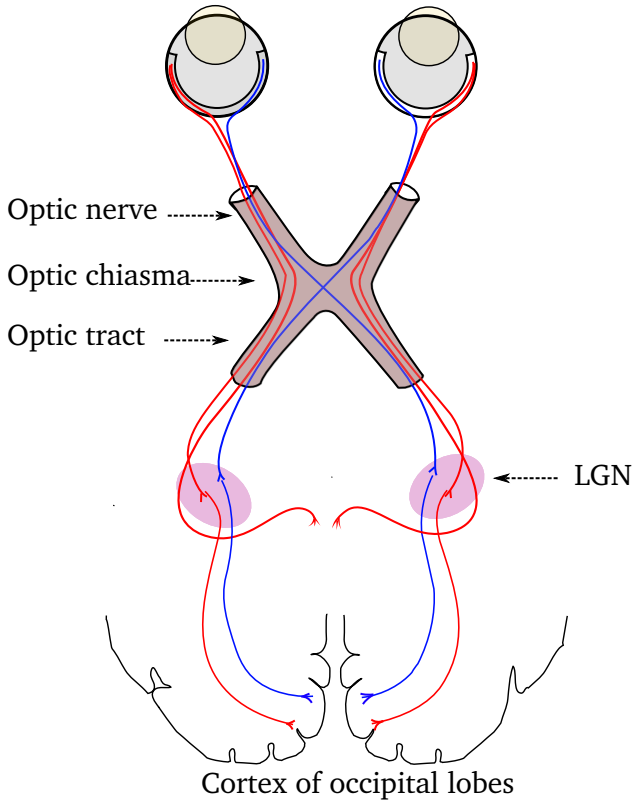


Figure 1.2: Due to the convexity of the eye lens, visual information on the right side of the scene is projected on the left side of both eyes and vice verse. The visual information is then transmitted through *optic nerves*, either temporally or nasally. Optic nerves meet at the *optic chiasm* located at the base of the hypothalamus, forming *optic tracts*. These tracts terminate at the LGN.

and object recognition are, in contrast, mediated largely by the ventral stream [4]. For example, identifying a cup as an object of interest (what) is accomplished through the ventral stream, whereas information about locating the cup (where), reaching for it, and grasping it (how) is provided by the dorsal stream. Furthermore, the dorsal stream uses an egocentric frame of reference computing the object properties relative to the observer while the ventral stream employs a scene-based frame of reference providing a detailed representation of the visual world. Our visual perception is dependent on both dorsal and ventral streams as shown in both behavioural [5] and biological studies [4]. Even though these streams are different pathways with different functionality, there is anatomical and behavioural evidence illustrating the cross talk between them at both

perceptual and physiological levels [4] [5].

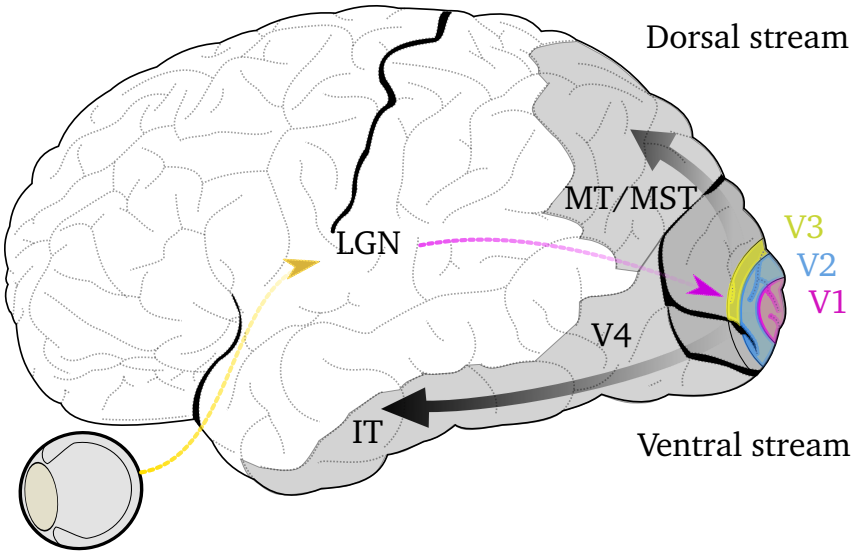


Figure 1.3: Dorsal-Ventral Streams: the dorsal stream runs forward from V1 into the posterior parietal lobe through V3, MT, and MST. The ventral stream stretches from V1 downward into the inferior temporal lobe through V4 and IT visual areas.

1.1.2 Visual Cues

The complexity of processing a visual scene increases as the information travels from the retina to the brain. Early stages of visual perception involves unconscious, rapid (200 msec - 250 msec [6]) accumulation of information about the visual scene provided by *pre-attentive features*. These features are often classified into four categories of *colour*, *form*, *movement*, and *spatial localization*[1] [7]. It is believed that the brain process these cues at the very low-level visual areas to have an initial estimate of the environment's properties and then combine them based on their statistical reliability.

[1] e.g., hue, intensity (colour) - line orientation, length, width, collinearity, size, curvature, spatial grouping, added marks, and numerosity (form) - direction of motion, flicker (movement) - 2D position, stereoscopic depth, concavity/convexity produced by shading (spatial localization).

Colour

Processing colours begins immediately after receiving light by a specialized type of retinal neurons called cone photoreceptors. There are three types of cone cells, each sensitive to a particular wavelength of light[2]. The light's wavelength and energy carry information about the hue (chromatic) and the luminance (achromatic) characteristics respectively. Colour perception is a result of processing both achromatic and chromatic signals. Our visual system is extremely sensitive to both hue and luminance[3]. For instance, identifying the red circle among the sea of blue ones (as shown in Figure 1.4) is a rapid, effortless process [6]. In this example, our brain unconsciously groups stimuli together if they share a similar hue (blue) and differentiate them if they don't (red vs. blue) [8]. Luminance also can serve as a visual cue to make a stimulus distinguishable. Areas with high contrast[4] not only attract the visual attention [9], but also play a key role in edge detection.

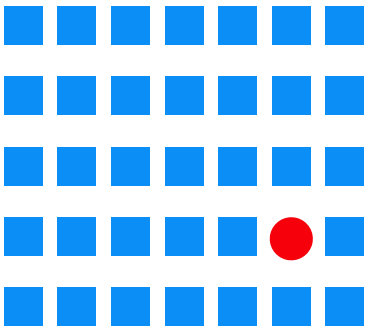


Figure 1.4: Detecting colour is a pre-attentive process; e.g., the red circle within the sea of blue squares

Form

If one picks up a pen and looks at it from different viewing angles, the perception of the pen's physical shape (i.e. distal stimulus) does not change. Meanwhile, the pattern

[2]S-cells for short-wavelength (blue), M-cells for medium-wavelength (green), and L-cells for long-wavelength (red) light

[3]The subjective impression of luminance is known as brightness

[4]Contrast is the difference in luminance and/or colour

of energy creating retinal images (i.e. proximal stimulus) is changing and so activating different sets of receptors. Despite the difference between the distal and the proximal stimuli, our perception of the pen remains constant; the pen is perceived as having the same shape regardless of the viewing angle, illumination, and its location in the visual field. This *perceptual constancy* in *object recognition* allows us to recognize and identify certain objects despite time-varying changes in the received signal. Object recognition involves a hierarchy of processes that begins with low-level recognition of *forms*, that is dots, lines, contours, and shapes. Recognition of contours, in particular, helps us to recognize shape, orientation, and relative depth, especially when other visual cues are lacking. This information is then integrated with other information collected from motion, colour, and depth providing information on the overall form. The result is then matched with structural descriptions in memory. Top-down processes such as familiarity and recollection may also provide information about the object of interest and its surrounding environment. Similar to colour perception, the ventral stream appears to be responsible for analysis of form and object recognition.

Perhaps one of the most notable theories in object recognition is David Marr's theory of vision [10]. According to his theory, object recognition involves three major stages of *primal sketch*, $2^1/_2$*D sketch*, and *3D model*. In primal sketch, low-level visual features such as edges, contours, and regions are extracted to explain local 2D geometries in an image. The physical information about material properties (e.g., colour, texture, shininess) and egocentric properties of visible surfaces (e.g., depth, orientation) are collected during the $2^1/_2$D sketch via illuminant, surface reflectance, surface shape, and the vantage point of observer. Information collected in the first two stages might suffice for proper perception. In some case, however, the spatial locations of objects are processed and related to their surrounding environment for 3D representation and more abstract definitions of the visual scene.

Motion

Motion is perceived when the light emitted from a moving object transverses the retina and fires neurons in a sequential pattern [4] (Figure 1.5-Left). Pursuing movement of a moving object also causes the perception of motion. In this case, the same part of retina is stimulated and so the information about the moving object comes from the movement of the eyes/head (Figure 1.5-Right) [4]. There is an exception to this phenomenon, however: In *beta-* and *phi phenomena*, the *apparent motion* is caused not by changes in the retinal images or movement of the head/eyes but because of rapid sequential changes in the static image giving an illusion of motion[5]. In the beta phenomenon, changes of static images occur with a speed faster than that which human eyes can resolve, causing the illusion of motion (despite the fact that there is no movement per se). In the phi phenomenon, changes in luminance of a stationary image causes the sensation of movement. In both cases, the visual system detects rapid changes in luminance at a certain point on the retina and correlates that with the luminance of its neighbours. In humans, motion perception, and the integration of local motions into a global percepts, largely take place in MT (V5) as part of the dorsal stream.

Depth cues

Depth perception involves consolidation of different monocular and binocular cues to form a three-dimensional image of the world from the two-dimensional images projected on our retina. Binocular cues, such as *stereopsis* provide information about depth when a scene is observed with both eyes. *Stereopsis* is the process of perceiving depth from two slightly different images of the world projected on the left and right retinas. The horizontal difference between these two retinal images, known as binocular disparity, is used by disparity-tuned cells in several cortical visual areas (such as V2

[5]The apparent motion is the basis of animation and motion pictures.

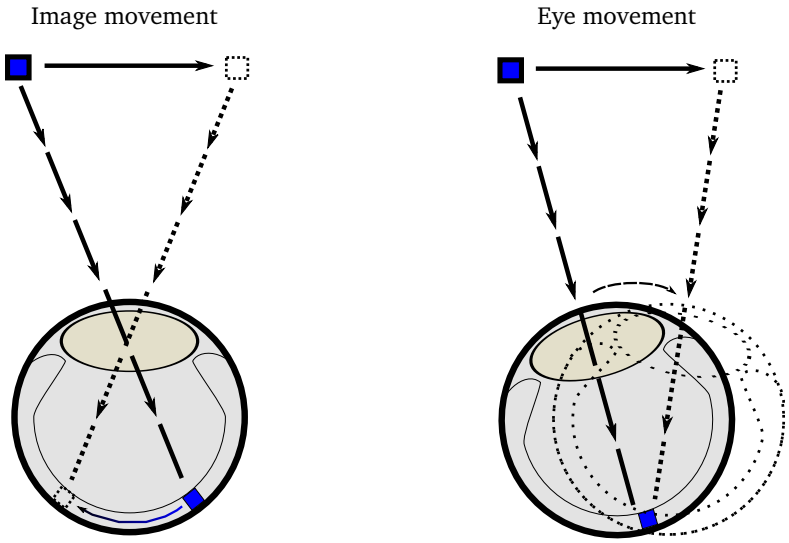


Figure 1.5: Information about the motion can be perceived by either following or holding the eyes still in front of a moving object (inspired from [4])

and V3) to compute depth. *Stereoacuity* is the acuteness of stereopsis representing the sensitivity to small disparity differences. Luminance, spatial frequency, observation time, and contours can influence stereacuity. For instance, at a low luminance level, stereoscopic acuity declines dramatically while increases in spatial frequency or fixation time improves it. Stereopsis is the most important binocular cue for interpreting short distances (Figure 1.6). Stereopsis is discussed in more detail in chapter 2.

Vergence is another binocular cue providing depth information for distances less than 10 meters (Figure 1.6). When one begins to fixate on a near object, this inward movement of the eyes (*convergence*) stretches the extraocular muscles, which in turn, sends signals to the visual cortex to be used for interpreting depth. A similar process occurs when the eyes rotate outwards to focus on more distant objects (*divergence*). Unlike binocular cues, monocular cues can be detected with one eye only to perceive depth. One of the most effective monocular cues is *motion parallax*, defined as the discrepancy in perceived speed as a result of differences in distance, i.e. when objects move at the same objective speed relative to the observer, closer objects traverse the retina faster than further objects. The speed by which the objects move as well as

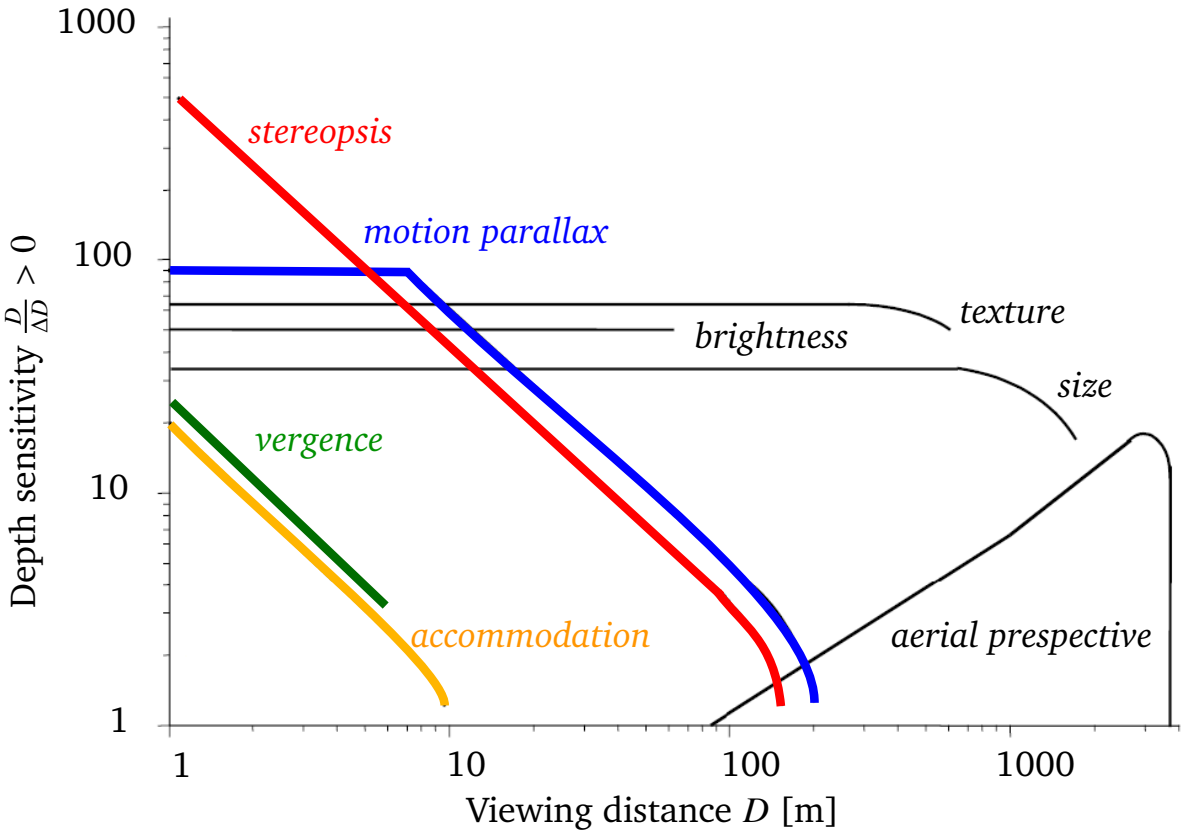


Figure 1.6: The sensitivity of depth cues as a function of viewing distance [11]

their direction of movement may provide sufficient information to perceive the relative distance between them (Figure 1.7). Interestingly, perception of depth in animals which lack developed binocular vision, is heavily dependent on motion parallax [12]. Motion parallax works best for short distances. Another effective cue, *occlusion*, happens when an object is partially occluded by another object within the same visual direction. Other monocular cues such as relative size, familiar size, perspective, aerial perspective, depth from motion, accommodation, texture gradient, and distribution of shadows and illumination can also provide egocentric and relative depth information. These cues are further detailed in Table 2.1.

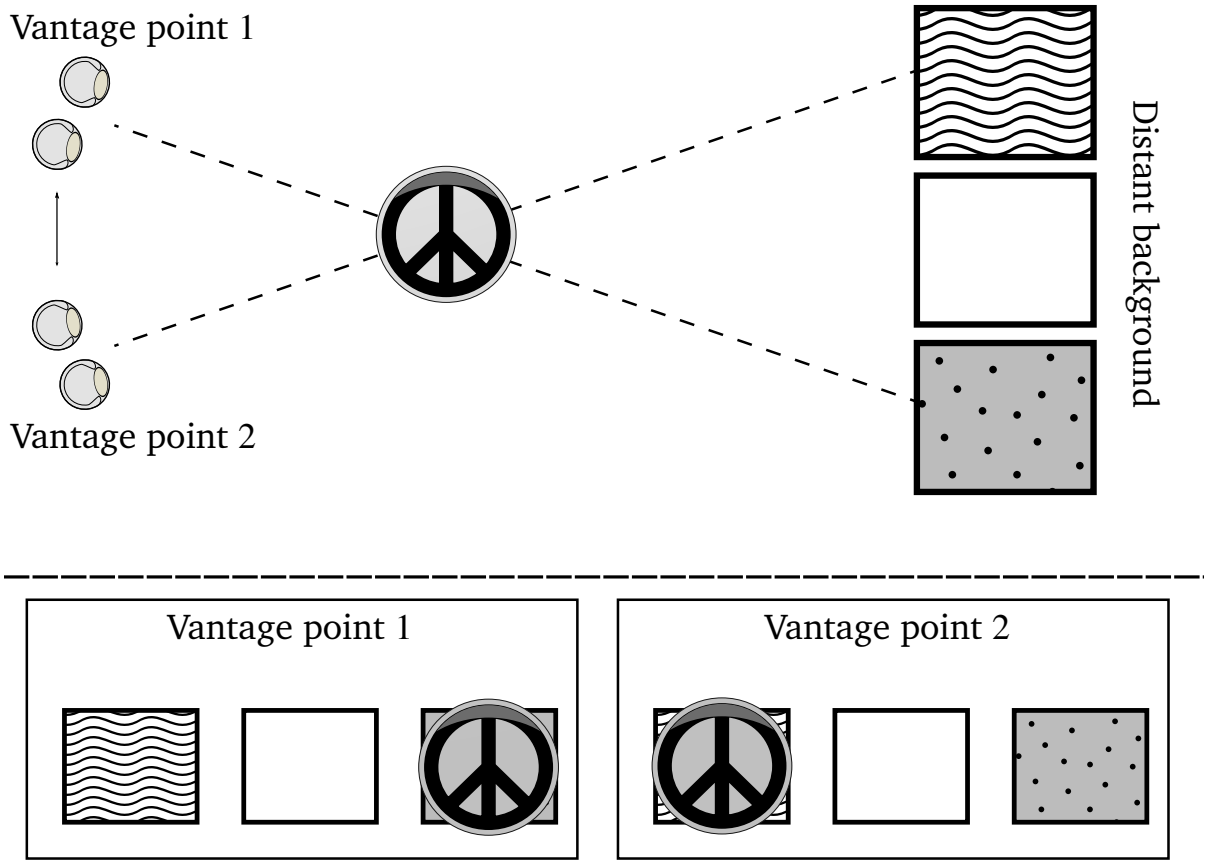


Figure 1.7: An example of motion Parallax

Translucency

Although translucency is not a visual cue, it is an important aspect of medical image visualization. Translucency allows for simultaneous visualization of multiple objects that would be impossible otherwise. Studies have shown that spatial reasoning (section 1.2.3) can be improved by concurrently presenting objects, as opposed to visually recalling the location of one with respect to another [13]. Additionally, translucency allows more information to be communicated in a shorter time, which in turn increases the efficiency of information transfer. The perception of translucency can be evoked by different factors such as refraction and world knowledge. However, *x-junctions* are perhaps the most important cues, influencing the perception of translucency. In x-junctions, the relative depth can be inferred from the order of luminance magnitude

along the contours of overlapping objects. As illustrated in Figure 1.8, in computer graphics, x-junctions with different patterns of luminance may provide ambiguous, unambiguous, and inconsistent information about the depth of overlapping objects [14].

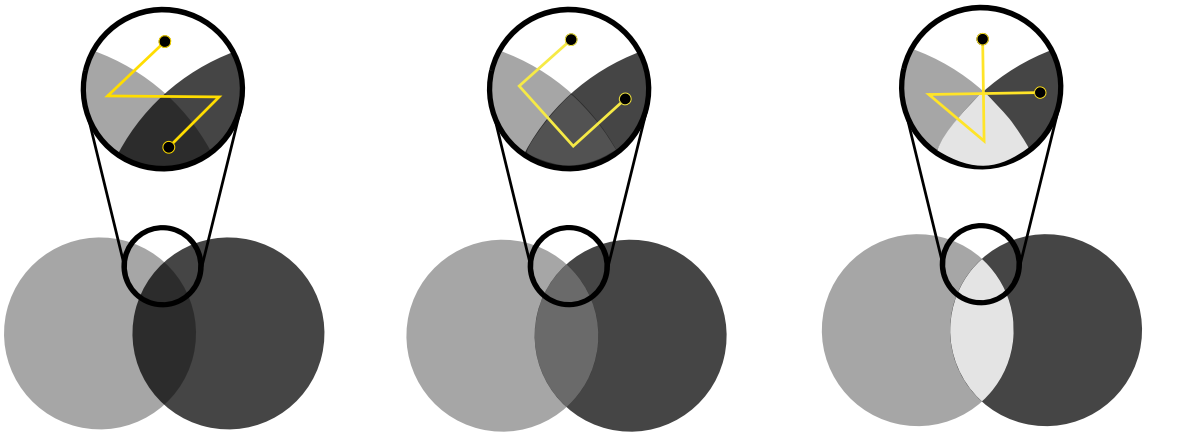


Figure 1.8: X-junction translucency: The ambiguity of the scene is determined by the order of luminance (Yellow lines correspond to the order of luminance from the brightest to the darkest). [14]

1.1.3 Bottom-Up & Top-Down Processes

Visual perception is a dynamic interaction between *bottom-up* and *top-down* processes working together and simultaneously, with feedforward and feedback interactions between all layers of our visual system. The bottom-up process is driven by sensory information from the physical world. The top-down process is driven by our knowledge, beliefs, expectations, and intentions. Both bottom-up and top-down processing streams interact at every level of visual perception to produce the best interpretation of the surrounding environment [15].

Bottom-up processing

The previous section discussed visual cues which are largely derived from bottom-up processing. Bottom-up processing of visual information is performed quickly, uncon-

sciously, and in parallel, with no interference of our goal-driven, endogenous attention. The task of the bottom-up processing is to detect and locate pre-attentive features within the entire visual field, providing information about the scene for the subsequent processes.

Top-down processing

As bottom-up information flows in from our sensory inputs, top-down information affects earlier processes in accordance with our prior knowledge and expectations. The top-down processing uses the information stored in memory to complement the bottom-up process making the perception more accurate and relevant. Figure 1.9 illustrates how top-down processes can influence our perception: despite the ambiguity of the middle character, one is able to perceive the same object differently given through their prior knowledge of the English language. According to the classical theories, perception is a processing hierarchy in which top-down is associated with feedback connections while bottom-up denotes feedforward information flow. From this point of view, features are first combined into an object and then the object is matched to some representations in memory. Recent studies in neuroscience, however, shows that top-down processing is sometimes activated earlier than some lower-level bottom-up processes to promote efficient recognition. In these cases, the initial guess about an object or a scene is highly influenced by contextual associations between the stimulus and the scene [16]. In this approach, the number of object representations that must be considered is significantly reduced by the top-down processes. From an evolutionary standpoint, “such a rapid mechanism provides critical information when a quick response is necessary” [17].

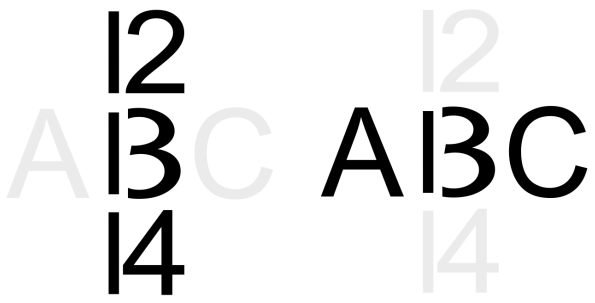


Figure 1.9: An example of top-down processing: the same character can be perceived either as letter 'B' or number '13'. This illustrates how our expectations can change our perception.

1.1.4 Measurement of Visual Perception: Psychophysics

"*What's the relationship between perceived depth and the neural activity in the primary visual cortex?*" To answer this question, it is not sufficient to study the behaviour of the brain's individual components (neural responses) but it is also necessary to interpret its overall behavioral activity (perceived depth) and correlate one to the other. Understanding the brain's behaviour, however, could be a daunting task because perception is an experience with temporal variations. This variation is due to the inherently probabilistic nature of sensory estimation, and that our senses are tuned to respond to a certain bandwidth of external stimuli. For example, if one estimates the depth of an object in a visual scene multiple times, the response often follows a Gaussian form probability density distribution with non-negligible variance [18]. To quantify this behaviour, *psychometric functions* (PFs) are acquired through conducting *psychophysical* experiments. PFs relate a perceptual property of a stimulus, often measured indirectly, with a corresponding physical property. In other words, PFs model the probability of brain's response to a certain property of the environment. In the simplest form, PFs are *sigmoid*-shape (Figure 1.10) with the y-axis representing the rate of performance (e.g. number of correct responses) and the x-axis representing the intensity of the stimulus (e.g. depth of an object). We can think of PFs as cumulative distributions of one's probabilistic behaviour in presence of certain stimuli.

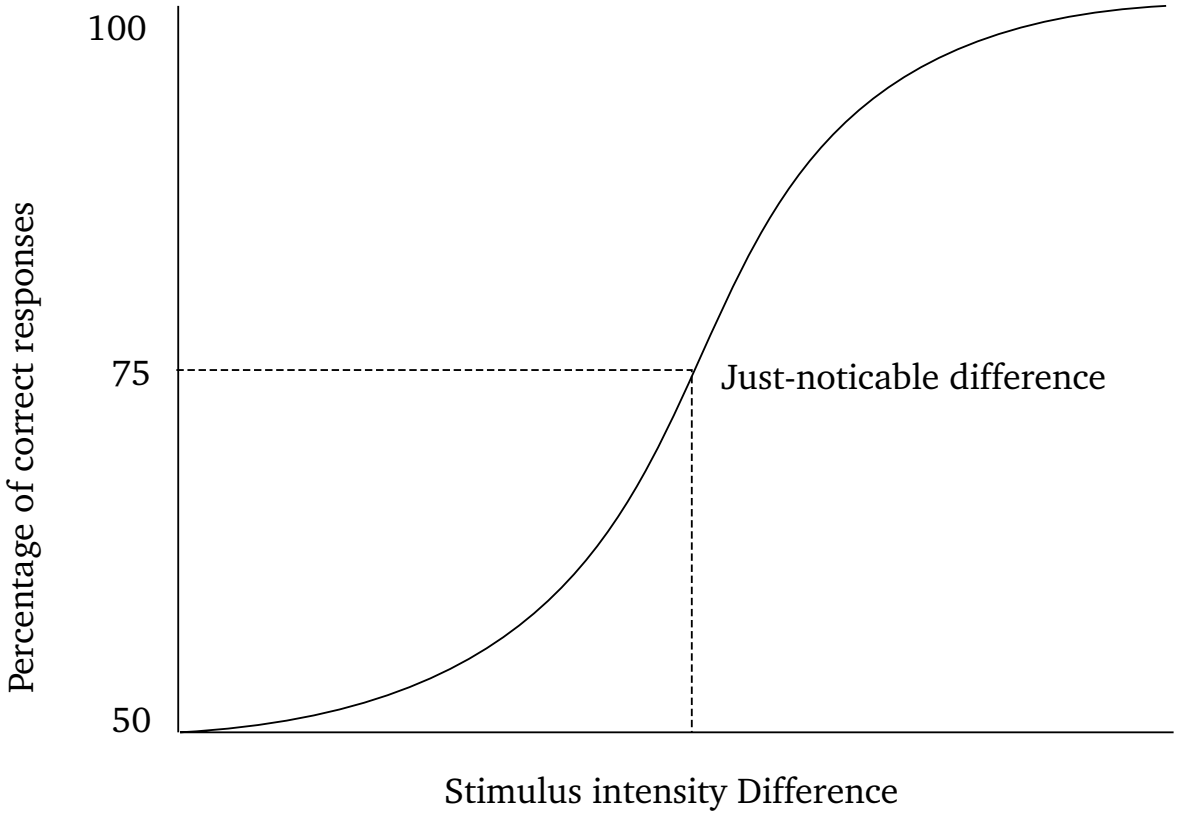


Figure 1.10: A typical psychometric function

Traditionally, psychophysical paradigms are carried out in laboratory conditions in which subjects are presented with a series of stimuli and asked to perform a certain task. In *magnitude estimation* experiments, for instance, subjects are required to subjectively rate the intensity of the presented stimuli on a given scale. *Matching* experiments involve matching the intensity of two stimuli by adjusting the intensity of one with respect to another. In *discrimination* experiments, subjects are asked to detect or discriminate small differences in the stimuli intensity. The design of discrimination experiments often follow one the following paradigms: *yes-no*, *two-alternative forced-choice* (2AFC), or *n-alternative forced-choice* (nAFC). Yes-no paradigms are often used to measure the *absolute threshold*, that is the level of intensity of a stimulus at which subjects are able to detect. In these experiments subjects are asked to confirm (yes) or refute (no) the existence of a stimulus at each trial. 2AFC (or nAFC) paradigms are used to measure the *differential threshold* (also known as *just-noticeable difference* or

JND), that is the smallest detectable difference between two (or n in case of nAFC) levels of a particular sensory stimulus (Figure 1.10). In 2AFC, two options are presented in every trial in a random order. One item of the pair is a *lure*, and the other one is a *target*. The observer's task is to report the observation that contains the target. Unlike yes-no paradigms, where subjects must report the presence or absence of a stimulus, forced-choice experiments are shown to be independent of the observers' subjective report of their personal perception[6].

1.1.5 Sensitivity Index

In yes-no paradigms the level of difficulty in detecting the target from background events can be measured as the proportion of targets which are correctly identified as such (i.e. (*hits*) / (*hits*+*miss*)), that is the *sensitivity* of the system. Sensitivity is therefore the degree of separation between targets and background events, or as described in the *signal detection theory*, between *signal* and *noise*. According to signal detection theory, the *sensitivity index* - commonly referred to as d' - is a measure of correlation between the level of task difficulty and the distance between signal and noise distributions. In yes-no experiments, d' is defined as $d' = (Z_{hit_{rate}} - Z_{FA_{rate}})$, where $Z_P(p \in [0\ 1])$ is the inverse of the cumulative distribution function. Similarly, in 2AFC experiments, d' provides a standard measure of performance as: $d' = (Z_{correct} - Z_{incorrect}) \times (\sqrt{2})^{-1}$. Naturally, higher d' values denote improved perceptual performance in locating the signal among noise. For more information, refer to [20, 21].

1.2 Cognition

Cognition refers to the conscious or unconscious mental processes that are involved in reasoning, problem solving, decision making, attention, memory, and perception (e.g.,

[6] It has been shown that individuals are not particularly good at reporting what they see or hear [19].

top-down processing mentioned in section 1.1.3). The effect of cognition in visualization research is rather profound. Cognitive processes not only derive the comprehension of visual information but also defines how to interact within environments.

1.2.1 Visual Attention

At any given moment, effectively processing the entire information contained in a natural scene is far beyond the brain's capacity. To cope with this potential *information overload*, the brain is equipped with *visual attention*, that is the cognitive mechanism of selecting behaviourally relevant information from the environment (and/or inhibiting irrelevant or interfering information) over space and over time [22]. This process may involve actively searching and processing visual information to locate the *target*. The efficiency and efficacy of such *visual search* can be increased by increasing the homogeneity of surrounding distractors [23], increasing the visual difference between the target and distractors (in features such as shape, size, orientation, motion, stereopsis, color, and lighting) [24], practice [25], and reducing the number of distractors.

1.2.2 Cognitive Load

Encoding, storing, and retrieving information is a complex cognitive process resulting from interaction between long term memory, short term memory, and working memory. *Working memory* is our conscious organizer and processor that passes information in small, incremental steps to *long term memory*, preventing random, rapid changes. Meanwhile, *short term memory* temporarily stores information to be later processed and organized by working memory [26]. Our working memory is limited in both content and retention capacities, particularly when dealing with novel information, and can be overloaded by too much information and/or too long of a retention. This phenomenon, known as *cognitive overload*, can severely disturb the efficiency and effec-

tiveness of information transfer [27]. There are three different kinds of cognitive load: *intrinsic*, *extraneous*, and *germane*. The intrinsic (or necessary) load is related to the difficulty of the content itself. The extraneous (or irrelevant) load is associated with unnecessary information such as distractions, and therefore should be reduced or eliminated. Germane (or relevant) load relates to the cognitive load *devoted* to processing new information into more advanced and complex structures known as *schema*. Interpretation of data, for a specific task at a specific level of expertise, poses an intrinsic load that cannot be changed. To facilitate the process of information transfer however, the intrinsic cognitive load needs to be efficiently managed without exceeding working memory limitation. This can be accomplished by providing adequate relevant prior knowledge while avoiding split-attention, redundant source of information, disappearance of information before being processed (transiency). As a rule of thumb, decreasing the extraneous load is necessary to improve the information transfer. It is important to note that reduction of cognitive load by lowering the difficulty of the problem is not always beneficial [28]. In the realm of training, for instance, mismatch between the level of expertise and problem difficulty can have negative effects on the quality of learning. This is even true when the inherent difficulty of the problem, the intrinsic load, is too low for trainees. Therefore, alignment of task difficulty and level of expertise is as important as lowering the extraneous load to enhance the learning process.

1.2.3 Visuospatial Reasoning

In 1974, Baddeley and Hitch proposed a model to describe how working memory functions [29]. Their multi-component model is composed of a *central executive* and its three slave systems, namely, *visuospatial sketchpad* (VS), *phonological loop*, and *episodic buffer* (Figure 1.11). The phonological loop stores verbal information, the VS deals with both visual and spatial information, and the episodic buffer holds integrated in-

formation across domains of visual, spatial, and phonological information [30]. The underlying cognitive processes, including attention, are supervised by the central executive system. Logie [31] has proposed that the VS can be further subdivided into a *visual cache*, which stores visual information about form and colour, and an *inner scribe*, which is a rehearsal mechanism for visual information and is responsible for spatial and movement information. The VS provides a temporary environment for *visuospatial reasoning*, that is our ability to perceive, visualize, simulate, manipulate, enact, and recall visual and spatial representations [32]; but perhaps the most important aspect of visuospatial reasoning is the *mental rotation* [33] and *translation* of visuospatial images. These transformations occur relative to three different classes of reference frames: *object-based* (relative to the object itslef), *egocentric-based* (relative to the self), and *environmental* (relative to fixed features of the environment) [34]. Egocentric-based transformations are either relative to the observer's perspective or to the end effectors (e.g. hands) [34]. While environmental transformations rarely occur, the object-based and egocentric-based transformations can be physically performed and/or *imagined*.

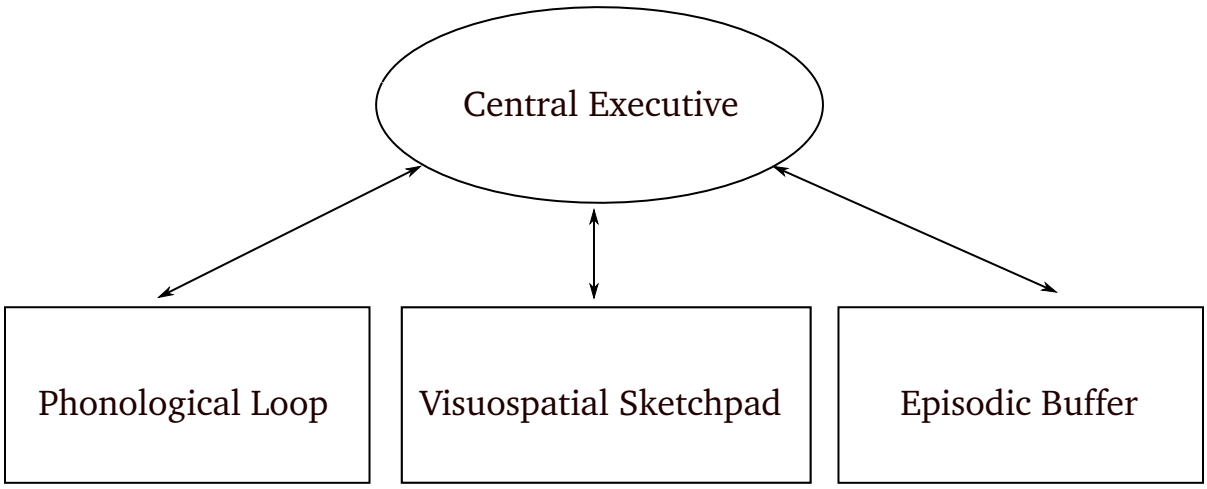


Figure 1.11: Multi-component model of working memory proposed by [29]

1.2.4 Measurement of Cognition

Measuring the cognitive load can be performed through performance-based and physiological measures, as well as subjective ratings [28]. *Performance-based* techniques often involve *dual task* experiments in which subjects' performance (i.e. time, accuracy) is measured with respect to a secondary stimulus while focusing on a primary task. Under this condition, the larger the cognitive load imposed by the primary task, the longer the reaction time and/or the lower the accuracy. *Physiological* measures involve measuring subjects' physiological changes such as heart rate and pupil size under the presence of cognitive load. These methods often suffer from low ecological validity - the degree to which the result of the study reflects the natural settings [28]. In *subjective ratings* method, individuals report their subjective experience of mental effort. Despite the body of evidence supporting the sensitivity of subjective ratings, one's subjective experience of difficulty may change significantly depending on his/her motivational and emotional condition [28]. Among available techniques for subjective workload measurement, the NASA task load index - NASA-TLX - is thought to be one of the most validated tools in this domain [35]. NASA-TLX is a post-hoc subjective multidimensional assessment tool developed by NASA after 40 laboratory simulations over a three year development cycle. This method involves two parts: first, subjects rate their experience of *mental demand*, *physical demand*, *temporal demand*, *performance*, *effort*, and *frustration* during the experiment. In the second part, subjects perform a pairwise comparison between these factors, selecting the one more relevant to the workload. The result of such comparison is then used to weight the initial ratings. The overall *task load index* is the average of weighted workload measures. NASA-TLX can be tailored to best suit the objective of the experiment (e.g., driving [36], surgery [37]).

1.3 Visuomotor Processing

Visuomotor skills refers to the coordination of muscular movements and vision to perform acts involving perception and action. In surgery, *eye-hand coordination* is a critical aspect of visuomotor processing in which visual information along with proprioception provide input to guide hand movements. Visuomotor skills - at the lowest level of action - are either *position-based*, that is a sequence of changes in position and orientation, or *selection-based*, that is making contact with, or grasping, an object. To perform these low-level motions, the cognitive system first decomposes the overarching goal into a sequence of tasks, which can be further subdivided into a number of sub-tasks. It is at the lowest level, that performance can be measured through behavioural studies. This is often accomplished in terms of the *speed* and *accuracy*. In this framework for performance, there is always a trade-off between speed and accuracy as some tasks favour speed over accuracy and some others favour accuracy over speed. Therefore, any methodology measuring a sub-task performance must always address both the speed and accuracy while combining the two in a way that respects the trade-off [7].

1.4 Layout of Dissertation

This thesis addresses psychophysical and psychological implications of visualization, interaction, and simulation in four different image-guided procedures with the goal of improving the quality of patient care. Chronologically speaking, the empirical findings offered in each chapter rationalize the approach taken in the following chapter.

[7] One of the pioneers of this field was Paul Fitts who developed a formula - commonly known as Fitts' law - to model the act of pointing [38]. Interestingly, Fitts' law can be altered and applied to any task with speed-accuracy trade-off, from low-level physical actions such as steering [39] to making perceptual judgment about an action that is observed but not performed [40].

1.4.1 Chapter 2: The Role of Stereopsis in Endoscopic Third Ventriculostomy

Endoscopic third ventriculostomy (ETV) is a minimally-invasive surgery to treat obstructive hydrocephalus by perforating the floor of the third ventricle. Although ETV is safer than the alternative (i.e. shunt placement), the risk of morbidity due to injury to the basilar artery increases as the floor of the third ventricle thickens and becomes opaque. In hydrocephalus patients, as the pressure from the cerebrospinal fluid increases, the ventricle expands and the floor deforms, warping around the underlying structures. Surface curvature could therefore be used as a landmark to locate and avoid the artery. Conventional endoscopes, however, provide an impoverished view of the surgical scene, impeding the detection of these structures. In Chapter 2, we establish a methodology to determine whether stereoendoscopy - as opposed to the conventional monocular approaches- can improve the perception of these structures, lowering the risk of ETV interventions.

1.4.2 Chapter 3: Visual Enhancement of MR Angiography Images to Facilitate Planning of Arteriovenous Malformation Interventions

Arteriovenous malformations (AVMs) are vascular lesions in which blood flows directly from the arterial to the venous system bypassing the normally interposed capillary bed. Treatment may involve surgical resection in which a surgeon examines preoperative images in order to gain a full understanding of the lesion's anatomy, its location, and its spatial relation with surrounding structures. Accordingly, in Chapter 3, I investigate the usability of a non-photorealistic contour enhancement technique, volume rendering, and stereopsis to improve upon identification and localization of vascular structures in non contrast-enhanced magnetic resonance angiography images. The results

of Chapter 2 on stereopsis is further extended in this chapter.

1.4.3 Chapter 4: Training for Planning Tumour Resection: Augmented Reality and Human Factors

Planning surgical interventions is a complex task, demanding a high degree of perceptual, cognitive, and sensorimotor skills to reduce intra- and postoperative complications. This process requires spatial reasoning to coordinate between the preoperatively acquired medical images and patient reference frames. In the case of neurosurgical interventions, traditional approaches to planning tend to focus on providing a means for visualizing medical images, but rarely support transformation between different spatial reference frames. Thus, surgeons often rely on their previous experience and intuition as their sole guide to perform mental transformation. In the case of junior residents, this may lead to longer operation times and increased chance of error under additional cognitive demands. In Chapter 4, I introduce a mixed augmented/virtual reality system to facilitate training for planning a common neurosurgical procedure, brain tumour resection. The proposed system is designed based on our previous findings on the effect of stereopsis and occlusion on surgical performance and evaluated with human factors explicitly in mind, alleviating the difficulty of mental transformation.

1.4.4 Chapter 5: A Complete Simulation Environment for Vertebroplasty Procedure

In addition to visual perception and its implications for image-guided interventions, the effect of cognition in surgical training is rather profound. Chapter 5 presents a unique simulation approach to vertebroplasty procedures, evaluating both surgeons' technical and non-technical cognitive skills during their immersion in a complete medical simulation environment. Vertebroplasty is a percutaneous procedure in which bone cement

is injected into a fractured vertebra in order to restore its stability. Two crisis scenarios are included for the evaluation of the effect of interruption on the low-level surgical skills. Validation and evaluation of our work is conducted together with 19 junior surgeons in order to provide a qualitative measure of usability, assess vertebroplasty technical performance of the surgeon, and associate between mental workload and surgical performance during crises. Training these mixed-mode scenarios can thereby be evaluated on our platform, allowing for improved assessment and a stronger foundation for credentialing, with the potential to reduce the occurrence of adverse events in the OR[8].

Closing remarks as well as future directions are discussed in Chapter 6.

[8]Chapter 5 is a result of collaboration with *Navigated Augmented Reality Visualization Systems* (NARVIS) *laboratory* at Technical University of Munich (TUM), Germany.

Bibliography

[1] Kevin Cleary and Terry M Peters. Image-guided interventions: technology review and clinical applications. *Annual review of biomedical engineering*, 12:119–142, 2010.

[2] T. Peters and K. Cleary. *Image-Guided Interventions: Technology and Applications*. SpringerLink Engineering. Springer, 2008.

[3] Margaret Livingstone and David Hubel. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, 240(4853):740–749, 1988.

[4] Robert H Wurtz and Erik R Kandel. Perception of motion, depth and form. *Principles of neural science*, pages 548–571, 2000.

[5] Melvyn A Goodale and A David Milner. Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1):20–25, 1992.

[6] Christopher G Healey. Perceptual techniques for scientific visualization. In *SIGGRAPH99 Course*. Citeseer, 1999.

[7] Colin Ware. *Information visualization: perception for design*. Elsevier, 2012.

[8] Maureen Stone. Choosing colors for data visualization. 2006.

[9] Pamela Reinagel and Anthony M Zador. Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems*, 10(4):341–350, 1999.

[10] David Marr, S Lal, and HB Barlow. Visual information processing: The structure and creation of visual representations [and discussion]. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 290(1038):199–218, 1980.

[11] Shojiro Nagata. How to reinforce perception of depth in single two-dimensional pictures. *Proceedings of the society for information display*, 25(3):239–246, 1984.

[12] Karl Kral. Behavioural–analytical studies of the role of head movements in depth perception in insects, birds and mammals. *Behavioural processes*, 64(1):1–12, 2003.

[13] Christina A Burbeck and Yen Lee Yap. Spatiotemporal limitations in bisection and separation discrimination. *Vision research*, 30(11):1573–1586, 1990.

[14] Roland W Fleming and Heinrich H Bülthoff. Low-level image cues in the perception of translucent materials. *ACM Transactions on Applied Perception (TAP)*, 2(3):346–382, 2005.

[15] RA Kinchla and JM Wolfe. The order of visual processing:top-down,bottom-up, or middle-out. *Perception & psychophysics*, 25(3):225–231, 1979.

[16] Mark J Fenske, Elissa Aminoff, Nurit Gronau, and Moshe Bar. Top-down facilitation of visual object recognition: object-based and context-based contributions. *Progress in brain research*, 155:3–21, 2006.

[17] Moshe Bar. A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, 15(4):600–609, 2003.

[18] Günther Knoblich. *Human body perception from the inside out*. Oxford University Press, 2006.

[19] Walter H Ehrenstein and Addie Ehrenstein. Psychophysical methods. In *Modern techniques in neuroscience research*, pages 1211–1241. Springer, 1999.

[20] Wilson P Tanner Jr and John A Swets. A decision-making theory of visual detection. *Psychological review*, 61(6):401, 1954.

[21] Micah R Bregman, Aniruddh D Patel, and Timothy Q Gentner. Stimulus-dependent flexibility in non-human auditory pitch processing. *Cognition*, 122(1):51–60, 2012.

[22] M Chun and J Wolfe. *Visual Attention, Blackwell Handbook of Perception*. Blackwell Handbooks of Experimental Psychology. Wiley, 2001.

[23] John Duncan. Boundary conditions on parallel processing in human vision. *Perception*, 18(4):457–469, 1989.

[24] H.E. Pashler. *Attention*. Studies in cognition. Psychology Press, 1998.

[25] Anne Treisman, Alfred Vieira, and Amy Hayes. Automaticity and preattentive processing. *The American journal of psychology*, pages 341–362, 1992.

[26] Jan L Plass, Roxana Moreno, and Roland Brünken. *Cognitive load theory*. Cambridge University Press, 2010.

[27] John Sweller, Paul Ayres, and Slava Kalyuga. *Cognitive load theory*, volume 1. Springer, 2011.

[28] Wolfgang Schnotz and Christian Kürschner. A reconsideration of cognitive load theory. *Educational Psychology Review*, 19(4):469–508, 2007.

[29] Alan D Baddeley and Graham J Hitch. Working memory. *The psychology of learning and motivation*, 8:47–89, 1974.

[30] Alan Baddeley. The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, 4(11):417–423, 2000.

[31] Irene C Mammarella, Francesca Pazzaglia, and Cesare Cornoldi. Evidence for different components in children's visuospatial working memory. *British Journal of Developmental Psychology*, 26(3):337–355, 2008.

[32] Daniel Reisberg. *The Oxford Handbook of Cognitive Psychology*. Oxford University Press, 2013.

[33] RN Shepard and Jacqueline Metzler. Mental rotation of three-dimensional objects. 1971.

[34] Jeffrey M Zacks and Pascale Michelon. Transformations of visuospatial images. *Behavioral and Cognitive Neuroscience Reviews*, 4(2):96–118, 2005.

[35] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183, 1988.

[36] A Pauzie. A method to assess the driver mental workload: the driving activity load index (dali). *IET Intelligent Transport Systems*, 2(4):315–322, 2008.

[37] Mark R Wilson, Jamie M Poolton, Neha Malhotra, Karen Ngo, Elizabeth Bright, and Rich SW Masters. Development and validation of a surgical workload measure: the surgery task load index (surg-tlx). *World journal of surgery*, 35(9):1961–1969, 2011.

[38] P. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47:381–391, 1954.

[39] Johnny Accot and Shumin Zhai. Performance evaluation of input devices in trajectory-based tasks: an application of the steering law. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 466–472. ACM, 1999.

[40] Marc Grosjean, Maggie Shiffrar, and Günther Knoblich. Fitts's law holds for action perception. *Psychological Science*, 18(2):95–99, 2007.

Chapter 2

The Role of Stereopsis in Endoscopic-Third Ventriculostomy

This chapter is adapted from '*Evaluation of a VR and Stereo-Endoscopic Tool to Facilitate 3rd Ventriculostomy*'[1].

My contribution to this chapter involved (i) designing phantoms, (ii) designing and conducting experiments, (iii) analyzing data, and (iv) writing manuscripts.

2.1 Introduction

2.1.1 Clinical Motivation: Endoscopic Third Ventriculostomy

Hydrocephalus, a neurological disorder that results in an abnormal accumulation of cerebrospinal fluid (CSF) within the brain, is one of the most common source of developmental disability among children, affecting one in every 1000 live births [1]. Hydrocephalus can be *communicating* or *obstructive*. Communicating hydrocephalus is

[1]Abhari, K., de Ribaupierre S., Peters T., Eagleson R. "Evaluation of a VR and Stereo-Endoscopic Tool to Facilitate 3rd Ventriculostomy", Studies in Health Technology and Informatics 163:1-7, (2011)

the result of insufficient reabsorption of CSF most often at the level of the arachnoid granulations, and can be caused by infection, subarachnoid hemorrhage, or be idiopathic. Obstructive hydrocephalus results from obstruction of CSF outflow along the ventricular system [1]. Potential etiologies of obstruction include mass lesions such as tumours, colloid cysts, but also aqueductal stenosis. In the US alone, obstructive hydrocephalus affects 375,000 people among the elderly population [1]. Obstructive hydrocephalus can be treated either with a shunt, draining CSF into other body cavities, or with a minimally-invasive surgery called Endoscopic Third Ventriculostomy (ETV). ETV interventions involve draining excessive CSF from the ventricular system into the interpeduncular cistern, creating a passage for CSF, by perforating the third ventricle and therefore bypassing the obstruction. In the current standard of care, *monocular endoscopes* are employed to navigate within the ventricular system, situate the region of interest, and create a stoma on the floor of the third ventricle.

Due to its lower rate of complications compared to shunt placement (e.g., shunt failure, infection, over-drainage, and slit-ventricle syndrome), in the last decade ETV has gradually become the procedure of choice for obstructive hydrocephalus. Furthermore, patients with multiple episodes of shunt malfunction can also be treated with ETV. However, with an overall complication rate of 8.5%, permanent morbidity rate of 2.4%, mortality rate of 0.21% [2], and the failure rate of 28% [3], patient safety is still of much concern. The most dreaded complication is perforating the basilar artery, which is located directly beneath the floor of the third ventricle, a few millimeters behind the clivus, supplying the brain and brainstem with blood (Figure 2.1). Rupture of the basilar artery would be considered catastrophic with loss of endoscopic visibility from bleeding, and could potentially result in significant stroke or death [2]. In the current standard of care, locating the basilar artery is only possible through direct endoscopic view of the operative field (Figure 2.2). This method, however, could be troublesome in those patients where an old infection, hemorrhage, or tumoral cells

wash off all the visible landmarks on the membrane (Figure 2.2). In these cases, because monocular endoscopes do not provide sufficient information and surgeons need to rely on their intuition and anatomical knowledge to locate and avoid the basilar artery, providing neurosurgeons with visual cues about the location of the artery can significantly lower the risk of ETV interventions and ultimately increase patient safety.

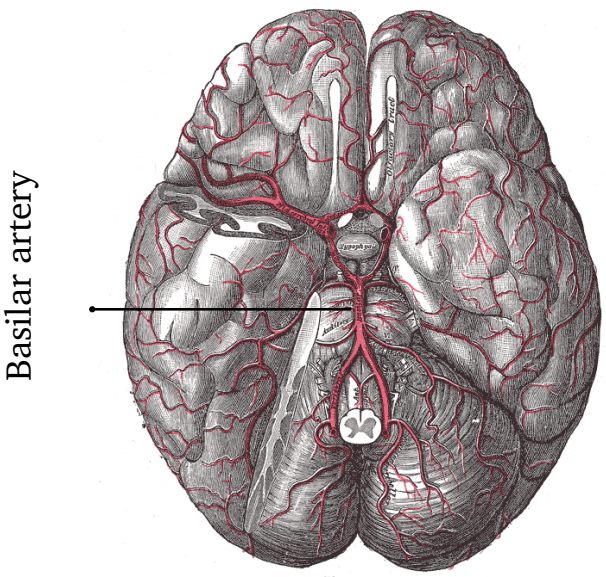


Figure 2.1: Anatomy of thrid ventricle: the basilar artery is located beneath the third ventricle, a few millimeters behind the clivus

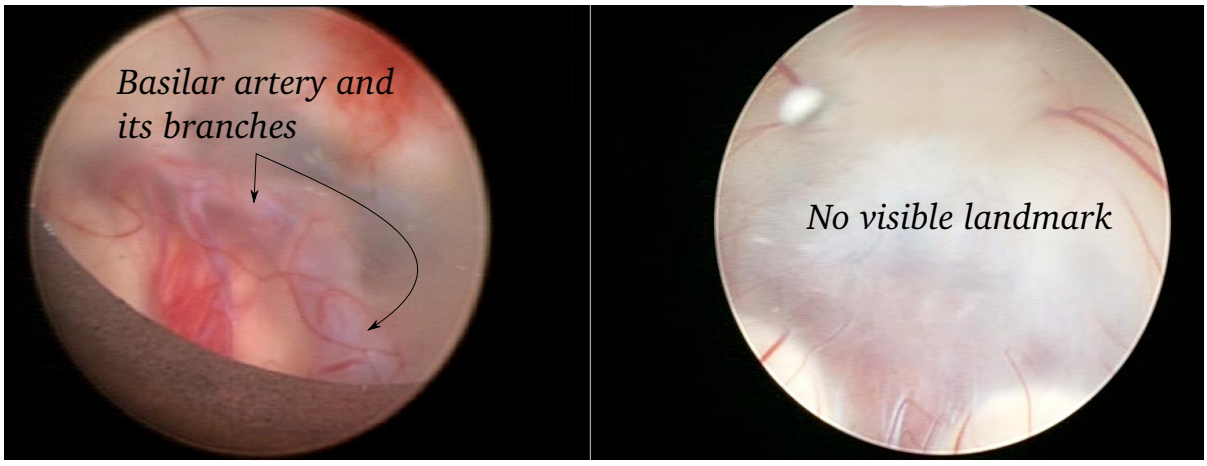


Figure 2.2: *Left*: An endoscopic view of the third ventricle during an ETV operation, *right*: lack of visible landmarks on the floor of the thrid ventricle when the tissue is thick and opaque

2.1.2 Hypothesis and Objectives

In hydrocephalic patients, the pressure of CSF inside the ventricular system gradually reshapes the structure of the ventricles. As a result, the floor of the third ventricle is pushed down, adopting the shape of the underlying structures. This mechanism, as illustrated in Figure 2.3, may lead to an irregular curvature on the floor of the third ventricle above the basilar artery. We believe that this curvature cannot be detected for most hydrocephalus patients unless additional information, such as stereopsis, is provided. We therefore hypothesize that stereoscopic endoscopes can provide necessary depth information to properly visualize critical structures and situate the target, preventing complications caused by injuring the basilar artery. Therefore, the overall objective of this study is to determine whether stereo-endoscopy can improve the localization of the basilar artery and also whether it can facilitate the process of perforation in hydrocephalus patients. In this study, we evaluate the feasibility of this approach using user studies conducted within both virtual and physical environments.

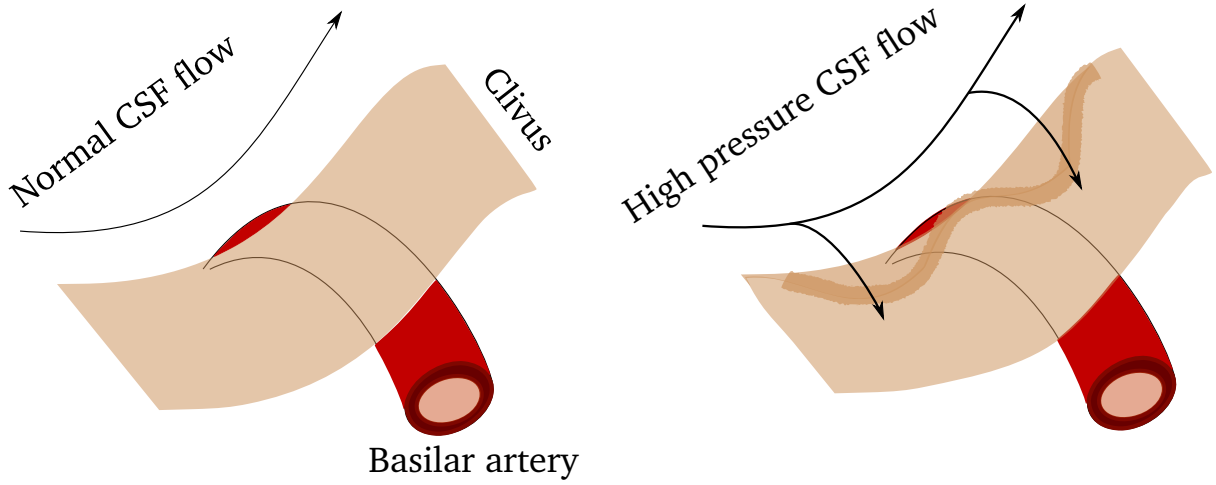


Figure 2.3: Because of the high pressure of CSF inside the ventricles, the bump above the basilar artery is more pronounced in hydrocephalus patients (*right*) compared to the normal population (*left*)

2.1.3 Background

Endoscopy in Neurosurgery

In 1910, an endoscope was employed for the first time to treat a neonate with hydrocephalus [4]. Since then, endoscopes have evolved dramatically and endoscopic procedures have become recognized methods of treatment in pituitary surgery [5], spinal procedures [6], and intracranial surgeries such as removal of intraventricular tumours or intracranial/intra-arachnoidal cysts [7]. Generally speaking, endoscopy is often the method of choice for neurosurgical interventions that deals with natural or pathologically formed cavities in the central nervous system. Endoscopes are also employed as an assistive technology in cerebrovascular aneurysms, microvascular decompression, and other microsurgical procedures to inspect the site of operation and locate structures out of line-of-sight, minimizing the drilling and retraction [7][8].

In the case of neurosurgery, endoscopes offer shorter operation time, less surgical trauma, and shorter hospital stays [9]. The main disadvantages of endoscopy, however, are restricted range of motion, degraded depth information, and the exhaustive amount of training needed to acquire adequate dexterity under their guidance [10].

Depth Perception in Neuroendoscopy

Depth cues - whether monocular or binocular - play an important part in surgical procedures, from which the distance between anatomical structures or between the end effectors (e.g., hands, instruments) and the operative field can be inferred. This is particularly important in certain endoscopic/laproscopic surgeries in which visual misperception is reported to be responsible for an astonishing 97% of surgical accidents (the other 3% was reported to be due to technical skills) [11]. In conventional endoscopy, however, monocular and binocular cues are either absent or severely degraded. Relying only on monocular cues leads to underestimation of depth (as shown in reaching

and grasping experiments [12]), and binocular cues are essentially absent in conventional neuroendoscopy. A list of depth cues and their role in neuroendoscopy is further detailed in Table 2.1.

Among the cues listed in Table 2.1, much attention has been given to stereopsis, as it is the only binocular cue that can be easily integrated into the surgical workflow, providing a strong sensation of depth. Stereoscopic images are captured by *stereoendoscopes* and displayed on passive or active 3D monitors. Standard monocular neuroendoscopes have a diameter of 3 to 6*mm*, suitable for navigating through the critical structures of the brain. Many stereoendoscopes with two cameras on the other hand, have a diameter of 8*mm* or greater, and thus afford no advantage over conventional microsurgery [7]. However, recent advances in technology have led to the development of stereoendoscopes with much smaller diameter (e.g., VisionSense iii with 4*mm* of diameter, VisionSense Corp., NY, USA) suitable for neurosurgical interventions. Despite these advancements, however, the use of stereoendoscopes in neurosurgery is still an open question, with many studies supporting or refuting their superiority over monocular endoscopes (section 2.1.3).

Table 2.1: List of depth cues and their role in endoscopy

Cue	Binocular or Monocular	Degraded	Absent	Description	Issue in endoscopy
Shadow	M	x	x	Shadows cast by anatomical structures can provide information about their shape (i.e. shape from shading) or depth.	In standard endoscopes, shadows are either absent when the light source coincides with the camera, or misleading when the location of the light source is unknown (Figure 2.4) [13, 14].
Perspective /Aerial perspective	M		x	Perspective and aerial perspective cues are the sensation of depth from, respectively, the convergence of parallel lines in the great distance and the decrease in the contrast/saturation of far away objects.	Due to the small size of the operative field, structures are imaged near to the camera, preventing perspective and aerial perspective cues to form.

Continued on Next Page...

Table 2.1 – Continued

Cue	Binocular or Monocular	Degraded	Absent	Description	Issue in endoscopy
Familiar Size	M	x	x	Familiar size cue is an estimation of depth by combining previous knowledge of an object with its projection on the retina.	Our estimation of the size of familiar (anatomical) structures -even within a familiar environment - can be inaccurate particularly in the absence of binocular cues [15].
Motion Parallax	M	x	x	Motion parallax is the apparent displacement of objects against their background, providing a strong sensation of relative depth.	In endoscopic environments, motion parallax is severely limited due to small range of motion of rigid endoscopes within the brain's cavities that prevents the relative movement between the camera and the structure of interest. Lateral head movement can also introduce motion parallax in natural environments, but in endoscopic interventions, it only confirms the flatness of the screen on which the endoscopic videos are projected on. This sense of *flatness* is even more accentuated by the frame around the screen, further degrading the sense of depth in such environments [13].
Occlusion	M			Occlusion refers to partial blockage of one object's view by another object.	Although occlusion may be present in endoscopic views, it only provides information about the *relative proximity* of objects as opposed to their relative depth.

Continued on Next Page. . .

Table 2.1 – Continued

Cue	Binocular or Monocular	Degraded	Absent	Description	Issue in endoscopy
Accommodation	M	x		Accommodation is the sensation of contracting or relaxing ciliary muscles to change the shape of the lens in order to focus on far or near objects, providing egocentric depth information.	In endoscopy, the eyes accommodate on the screen on which the endoscopic videos are displayed on. Therefore, accommodation provides depth information about the screen rather than the field of operation.
Vergence	B	x		Vergence is the kinesthetic sensations of the extraocular muscles when eyes are converging (or diverging) to focus on an object.	Vergence is ineffective in endoscopic settings because eyes converge to focus on the objects observed on the screen rather than the object itself. This necessitates to mentally rescale the view for correct visoumotor output.
Stereopsis	B		x	Stereopsis is the sensation of depth from two different images of the world projected on the left and right retinas.	Conventional endoscopes do not provide horizontal disparity, preventing the sensation of stereo.

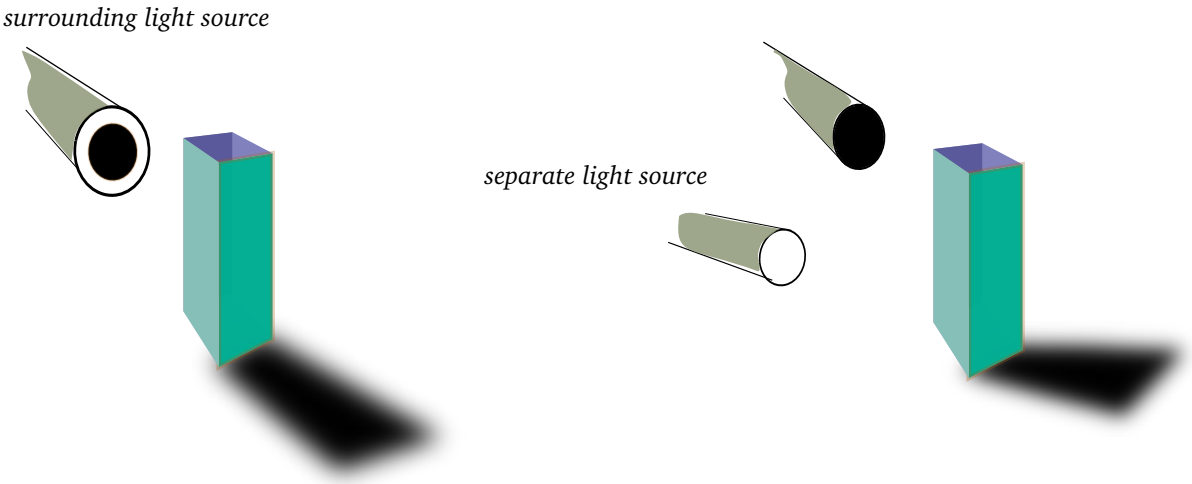


Figure 2.4: *Left*: Shadows cannot be seen in conventional endoscopy; *right*: In the images generated via shadow-forming endoscopes with a separate source of light, shadows are visible and can be used as a cue to perceive depth

Related Work

The clinical benefit of stereoendoscopy has been reported in sinus and skull base surgery [16, 17, 18], trans-sphenoidal surgery [19, 20, 21], and skull base reconstruction [22]. The results of non-clinical studies, however, are astonishingly mixed. Some studies have found that improved depth perception via stereoendoscopes leads to significantly faster performance and/or greater accuracy compared to 2D (monocular) endoscopes [23, 24, 25, 26, 27, 28], while others have found no significant difference between the two [29, 30, 31, 32, 33, 34], and in some cases subjects were shown to be even faster in 2D [35]. *Vergence-accommodation mismatch* has also been reported as a drawback of stereoendoscopy [36]: as illustrated in Figure 2.5, while the eyes constantly diverge and converge to look at different objects on the 3D screen, they always accommodate on the same surface. In open surgeries, similar to other natural settings, these two cues are always in synchrony, providing correct sensation of egocentric depth. The effects of stereoendoscopes and high-definition monocular endoscopes were investigated in a study done by Marcus et al. [37]. Their results revealed that while the high-definition monocular endoscopes led to greater accuracy, stereoendo-

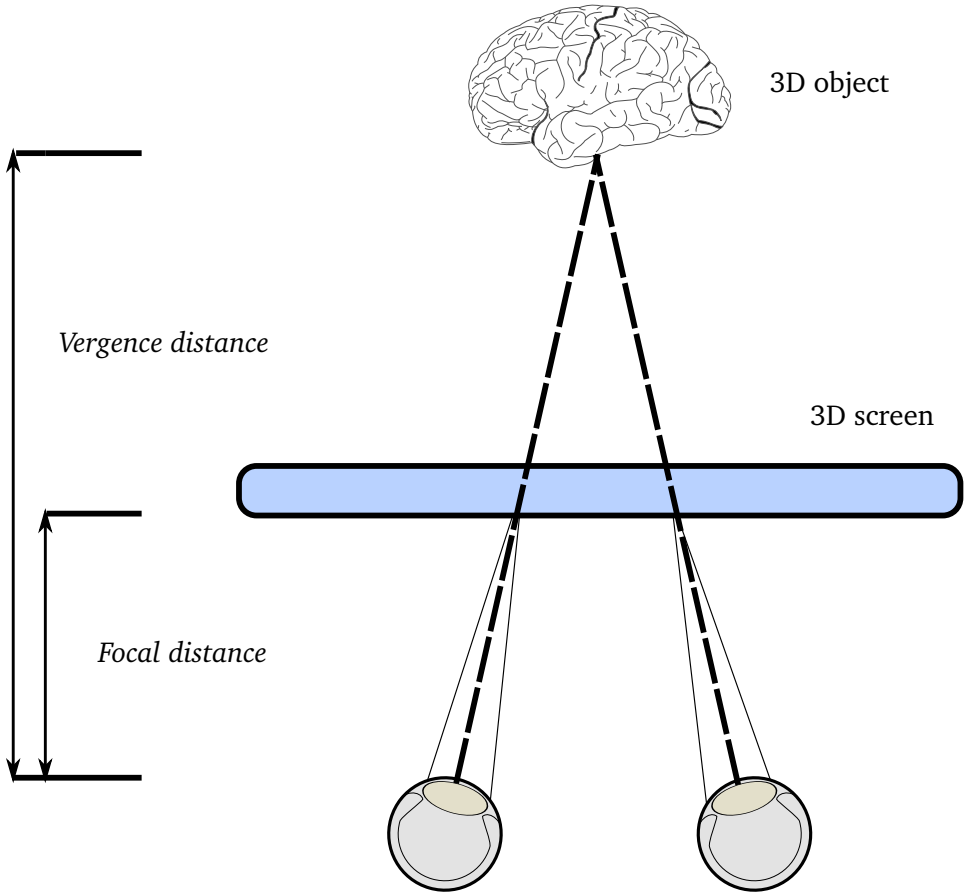


Figure 2.5: Vergence-accommodation mismatch in stereoendoscopy. In natural settings, these two cues are always in synchrony, providing correct sensation of egocentric depth.

scopes improved depth perception and shortened the completion time. The authors, therefore, concluded that both types of endoscopes have complementary effects on surgical performance, and that neither can fully compensate for the drawbacks of the other.

Stereopsis and Psychophysics

Stereoacuity, similar to visual acuity, is a measure of the perceptual capacity to detect small differences in depth using binocular vision. Stereoacuity can be measured through $2AFC$ experiments in which the binocular cue is the main source of depth information (refer to section 1.1.4 for details). In these experiments, the JND (just-

noticeable-difference) is measured as the difference in depth at which subjects correctly detect the *target* stimulus in 75% of the time (i.e. half-way between 50% or pure chance and 100% or perfect response). In this study, a *staircase* paradigm is employed to compute the JND. This paradigm requires fewer stimuli to reach the threshold compared to traditional psychophysical techniques. The intensity of stimuli increases and decreases systematically in order to detect the JND at the point where higher intensities could not be detected while lower intensities are easily detectable by subjects. To design an efficient staircase paradigm, it is suggested the following variables be estimated through pilot studies: where to *start* and *stop* the series, how large the *steps* are, and when to *modify* the series [38] (Figure 2.6). The staircase paradigm is most efficient when the intensity of the initial stimulus is chosen slightly larger than the JND; otherwise, many trials are required to reach the threshold, introducing fatigue and learning effects. As a rule of thumb, it can be set as the intensity at which level of performance reaches 95% [38]. In these experiments, step size is defined as the minimum difference between the intensities as we move from one trial to the next. Choosing a very small step size could be another source of inefficiency: if the steps are very small, many trials are required to reach the threshold. If the steps are very large, on the other hand, the JND would be insensitive to the changes in the stimuli since the intensity hovers around (without approaching) the threshold. Thus, an ideal step size is often defined as the intensity at which subjects perform discrimination halfway between chance and perfect performance, i.e. the estimated JND value. To maximize the efficiency, however, steps are often set to be slightly larger than JND in the beginning of the series and gradually become smaller when the final threshold value is about to be reached. Deciding when to end the series can be a compromise between a large series of stimuli for higher accuracy and a small number of trials for economy in time, minimizing the effect of fatigue. It is suggested that the end point can be set as the point where subjects' correct responses reach their plateau in the preliminary studies.

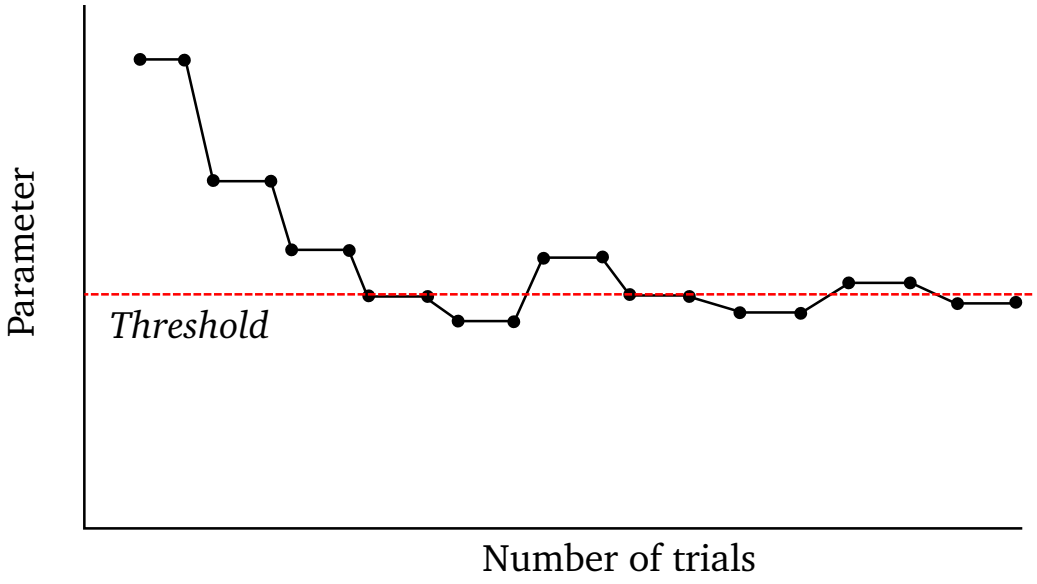


Figure 2.6: An example of a staircase paradigm

2.1.4 Contributions

The contributions of this study can be summarized as follows:

- Investigating the use of stereoendoscopes to detect critical structures in hydrocephalus patients;
- Exploring the role of stereopsis in avoiding the basilar artery and perforating through the third ventricle in ETV interventions.

2.2 Materials and Methods

To test our hypothesis, evaluation studies were conducted in two different phases. In phase 1, the use of stereopsis to identify the location of the basilar artery was examined. This is accomplished by comparing the ability of estimating the height of the surface above the artery in both mono- and stereoendoscopic views. In phase 2, a user study was conducted to examine the use of stereopsis to correctly situate the target location as a critical step in ETV interventions.

2.2.1 Phase 1

Phase 1 involved testing the ability of our subjects to distinguish between different structures in terms of their height in the presence or absence of stereopsis using 2AFC psychophysical experiments. Since the staircase paradigm was the method of choice, a preliminary study was conducted to estimate the predetermined factors (section 2.1.3) within a virtual environment that could be controlled with respect to the experimental variables. The results were then used to design a more realistic comparison experiment with the use of a stereoendoscope and physical phantoms.

Preliminary Study

In the preliminary study, several scale models were generated using a computer-aided design program (SolidWorks, Waltham, USA) based on the real anatomy and workspace geometry of the third ventricle (Figure 2.7). In these models, an endoscopic image of the floor of the third ventricle with no distinguishable monocular cues was mapped onto the surface with a slight elevation above the basilar artery, ranging in height from $0mm$ to $3mm$ with step value of $0.1mm$[2]. These models were visualized within the AtamaiViewer visualization environment (Atamai Inc., London, ON), presented on a FakeSpace™ display, a high-resolution stereoscopic system offering a large slanted table-top display for use in immersive visualization experiments (Figure 2.8). In each trial (total number of trials per subject: $N = 180$), subjects were asked to sit comfortably viewing the screen while wearing LCD shutter glasses, and presented with a pair of stimuli in mono or stereo (Figure 2.9). These models were placed side-by-side and observed from the top view angle, mimicking an endoscopic view in ETV operations. The task involved selecting the model with greatest (ventricular) protrusion by pressing the corresponding key on the keyboard. The order of stimuli, as well as the mode

[2]In this study, the height of the surface elevation above the artery is considered to be $\sim 1 - 3mm$, estimated by our collaborator expert neurosurgeons.

of visualization (stereo vs. mono), were counterbalanced to minimize the effects of learning and fatigue. The result of this study is presented in section 2.3.1.

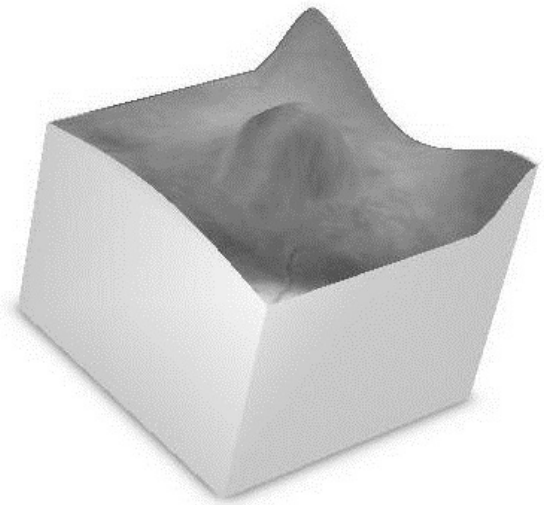


Figure 2.7: Model of the third ventricle

Comparison Experiment: Stereo vs. Mono

The VisionSense VSII 3D endoscope: is an FDA-approved, compact stereoendoscope (4*mm* in diameter, $\times 8$ magnification, Visionsense Corp., New York, USA), suitable for minimally-invasive neurosurgical procedures. The clinical applicability of VisionSense stereoendoscope has been demonstrated widely in the literature [17, 20, 16, 19, 21, 18, 22]. As illustrated in Figure 2.10, VisionSense is equipped with a lens in front with two pupil openings at the focal plane and a lenticular lens covering a CCD chip in the back. Under this arrangement, light rays pass through the center of each pupil and emerge as a parallel beam behind the lens. Since each lenticule covers exactly two pixel columns, rays are focused under the lenslets on the right and left side to generate right and left images. Such configuration allows for miniaturization of the endoscope. The VisionSense camera is used in this study along with physical phantoms to investigate the effect of stereopsis in detecting surface curvatures.

Figure 2.8: The environmental setup for the preliminary experiment

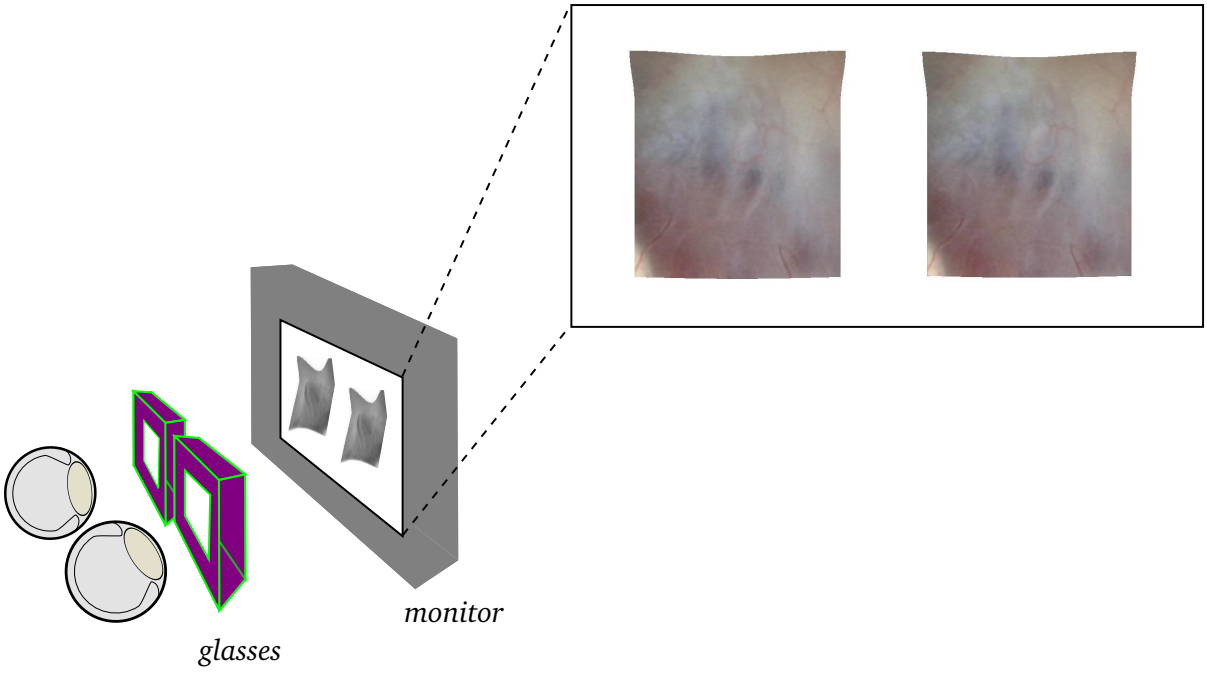


Figure 2.9: Two stimuli are placed side by side in 2AFC experiments

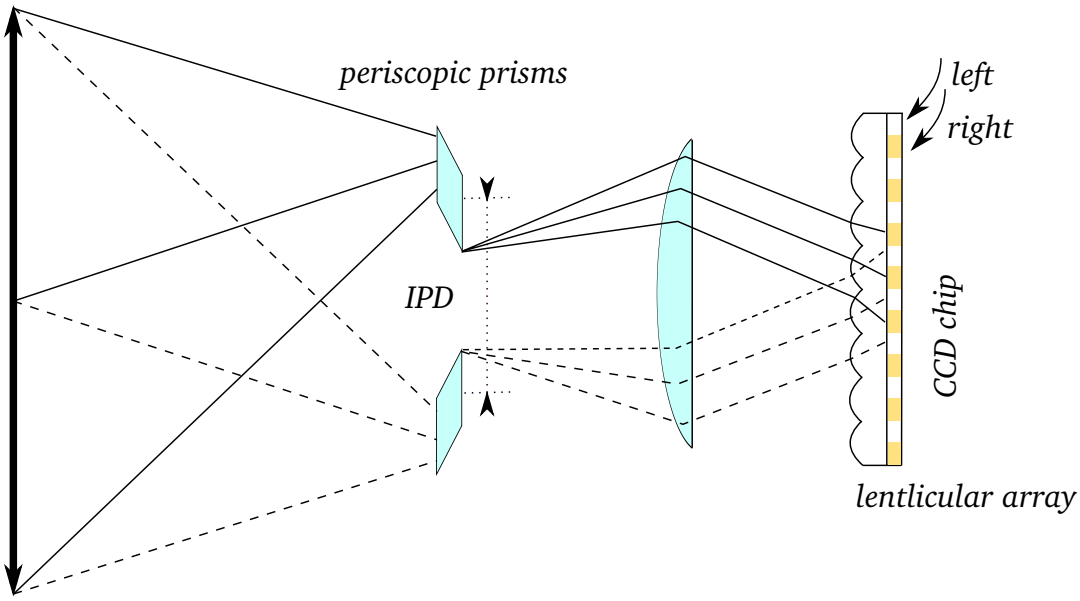


Figure 2.10: The underlying structure of the VisionSense camera: The rays passing through the two pupils emerge as parallel beams, reaching the left and right pixels on the CCD chip that is covered by a lenticular array. IPD corresponds to the distance between the two pupils.

Experiment: Similar to the virtual environment, each trial involved observing two side-by-side phantoms while wearing LCD shutter glasses (Figure 2.11). Subjects (n = 11) were asked to perform a discrimination task (in both mono and stereo) to select the phantom with a most prominent protrusion on the surface at every trial. The results are presented in section 2.3.1.

In order to make use of results obtained within the virtual environment, however, the VisionSense stereoendoscope and the virtual stereo environment were required to provide the same stereo effect (i.e. disparity). To fulfill this requirement, we varied the inter-ocular distance as well as the focal point of the virtual cameras. The distance between the blocks and the lens is also kept the same for both real and virtual environments.

The choice of phantom: Phantoms were first made using stereolithographic rapid-prototyping technology with $0.1mm$ of resolution based on our computer-generated models as seen in Figure 2.7. Unfortunately, this set-up resulted in some undesirable

effects including a gradient around the edges and a glare due to the reflection of the endoscope's light. Although these effects were not pronounced, they could potentially be used as monocular cues. Therefore, new phantoms with no pronounced monocular cues were produced from silicone with the resolution of $0.1mm$ (Figures 2.11 and 2.12).

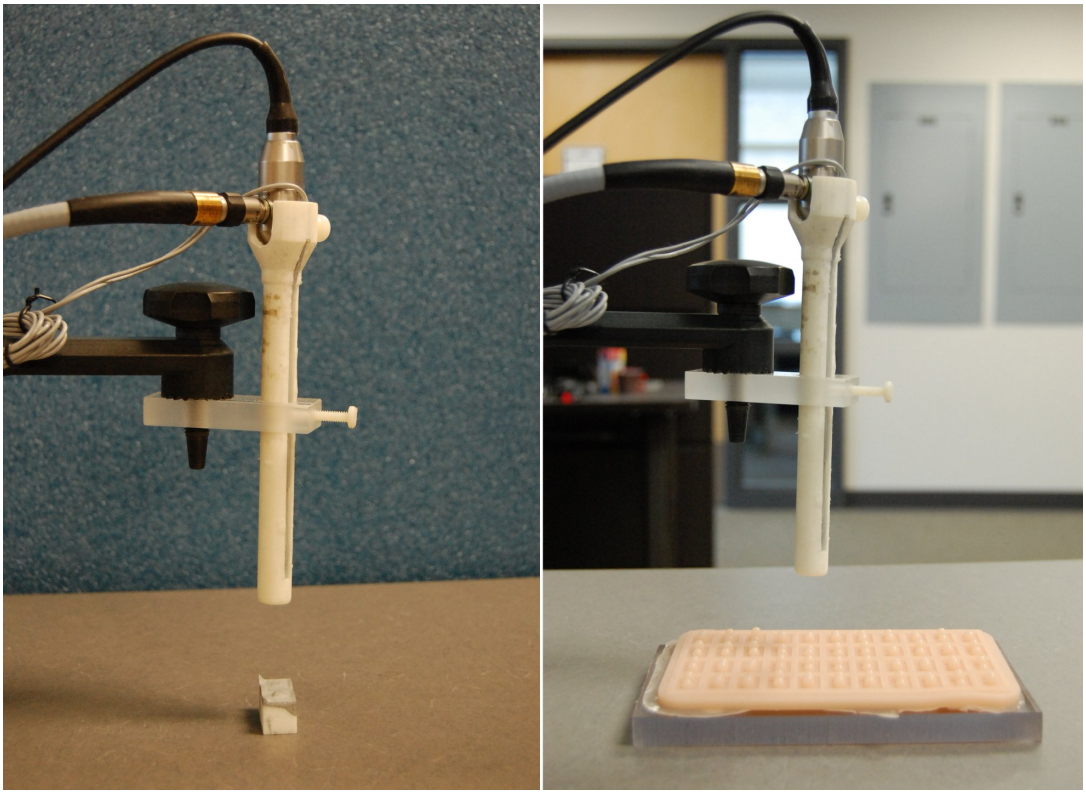


Figure 2.11: The environmental setup for the comparison experiment with the use of VisionSense camera

2.2.2 Phase 2

Phase 2 involved evaluating the performance with which a subject could localize a target that cannot be visually identified, but instead is estimated with respect to other locations. This approach is adopted because in ETV interventions, surgeons often pierce through the membrane at a point which does not have a corresponding visible landmark, instead, estimating its position relative to the basilar artery and the clivus. How-

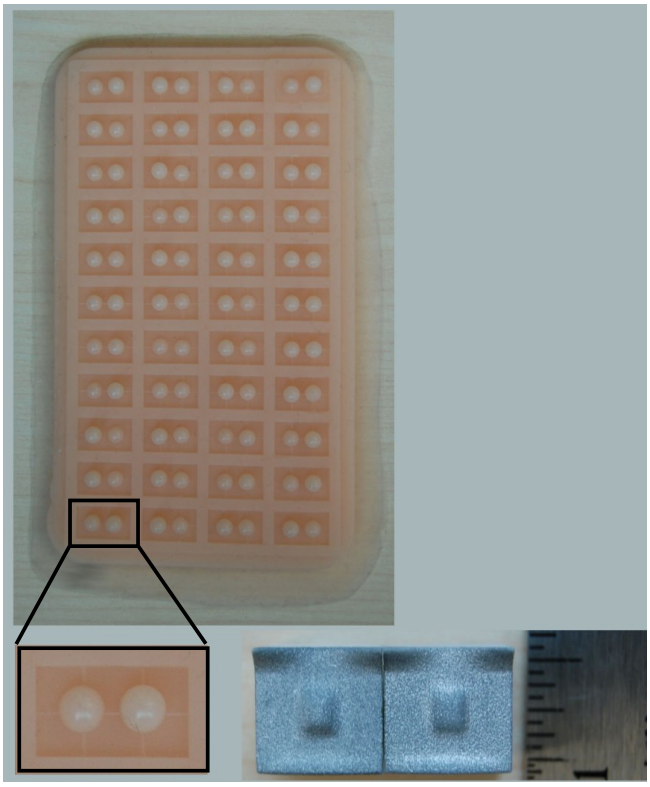


Figure 2.12: *bottom right*: First set of phantoms, *left*: Second set of phantoms made out of two-part silicon

ever, in many cases, each of these two structures can be imperceptible when viewed using a monocular endoscope. Even in the best viewing conditions, the structures are visible only through slight variations in the translucency of the clear membrane. Since the monocular cues are inadequate, targeting localization of the basilar artery and clivus may be improved using the stereo disparity cues that may be available from the small visible features in the texture of the membrane. Additionally, when targeting an unspecified location between two visible targets, one must consider the relative costs of making a targeting error. In our surgical context, if the basilar artery is touched accidentally, the result can be fatal. If the surgeon touches the membrane that covers the clivus, the monopole instrument cannot be pushed through at that point, and so a new target location must be selected. Consequently, the sense of optimal location is biased

towards the clivus[3]. Although there is a wide range of acceptable locations for the hole to be made, the monopole instrument is quite wide relative to the size of the free space. In considering all of these factors, we have established the following methodology for targeting performance in our experiments: Using computer-aided design software, a set of virtual models was programmed based on the real anatomy of the third ventricle and the distance between the basilar artery and the clivus. To simulate the cases where there is no visible cue on the surface of the third ventricle, a random-dot pattern was mapped onto the surface of our virtual models to eliminate distinguishable monocular cues (Figure 2.13). This type of texture does not provide required depth cues and therefore, the surface seems flat to an individual who observes the models monocularly from the top viewing angle (similar to what a surgeon might observe through a monocular endoscope in an ETV intervention). Under these circumstances, users could locate the target 50% of the time. Similar to the preliminary study in phase 1, subjects (n = 11: 1 expert, 10 novice graduate students) were asked to perform the task while sitting comfortably in front of the FakeSpace™ display and wearing LCD shutter glasses. Subjects went through a training session to learn about the anatomy of the third ventricle as well as the optimal location of the target. The task was defined as selecting a target location using a cursor that was controlled interactively, and resembled the tip of the monopolar instrument with $1mm$ of diameter (Figure 2.14). The targeting task was repeated for 150 trials per subject, over which the virtual height of the surface above the clivus and basilar artery were varied systematically over a range between $0mm$ and $3mm$ at $0.1mm$ increments. The location of these structures in the virtual endoscope view was also varied in their position at random locations within the visual field.

[3]This optimal location is estimated by our neurosurgeon collaborators as approximately about $^1/_3$ of the way in between the basilar artery and clivus (closer to the clivus)

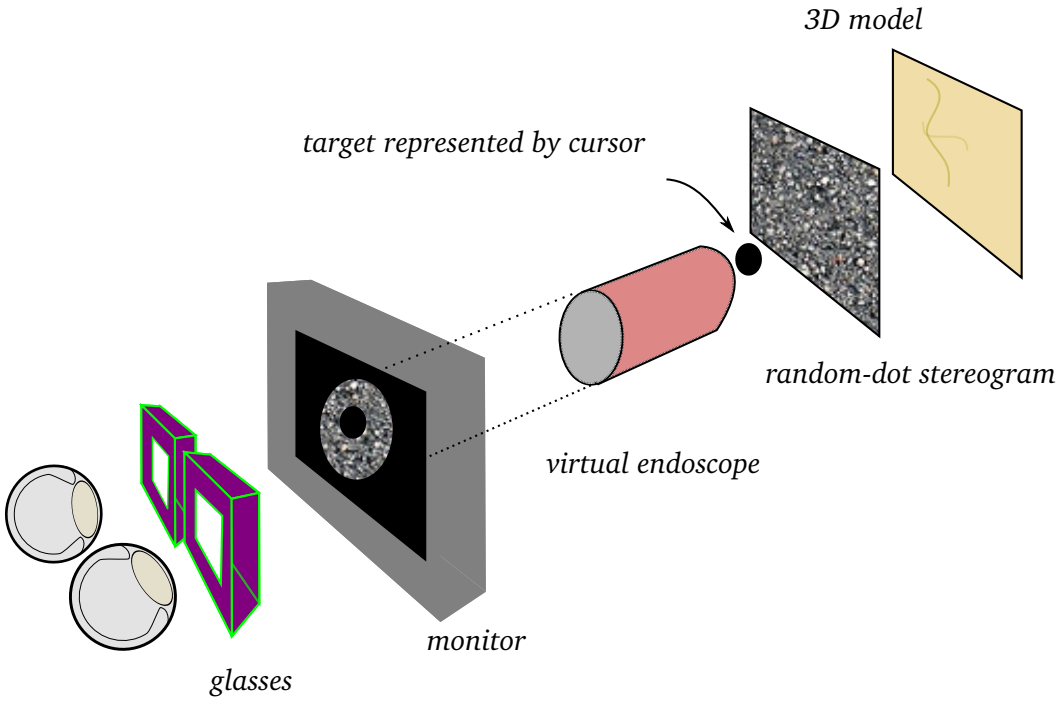


Figure 2.13: A random-dot pattern was mapped on the surface to eliminate the effect of monocular cues

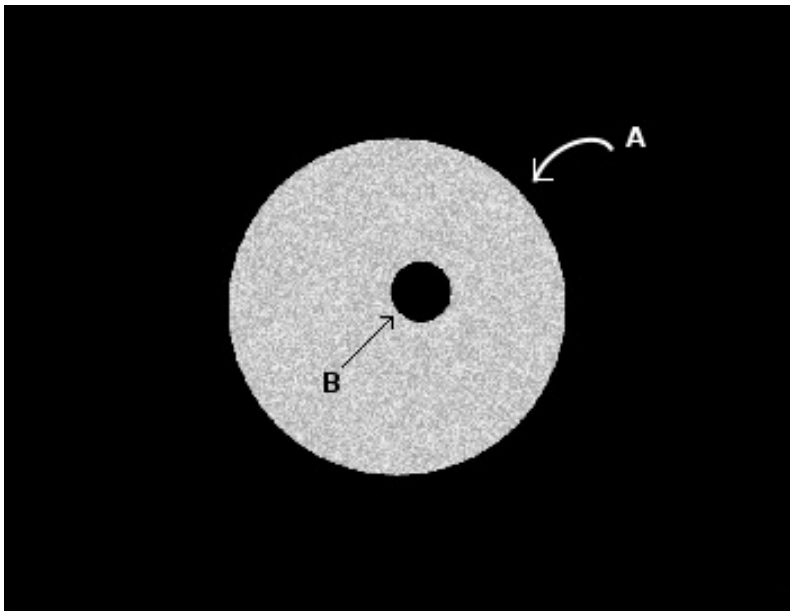


Figure 2.14: A) Virtual endoscopic view, B) Cursor represents the tip of the monopolar instrument

2.3 Results and Discussion

2.3.1 Phase 1: Results

Preliminary Study

As illustrated in Figure 2.15, the overall correct responses reached $\mu = 75\%$ and $\mu = 95\%$ when the height difference was approximately $0.5mm$ and $1.25mm$, respectively. Thus, $1.25mm$ was chosen as the initial intensity. For the step size, $0.5mm/2 = 0.25mm$ was chosen to increase the accuracy and maximize efficiency. Although after 40 trials, correct responses plateaued at $\mu \approx 100\%$, a larger number of trials was considered for the comparison experiment to compensate for lowering the step value.

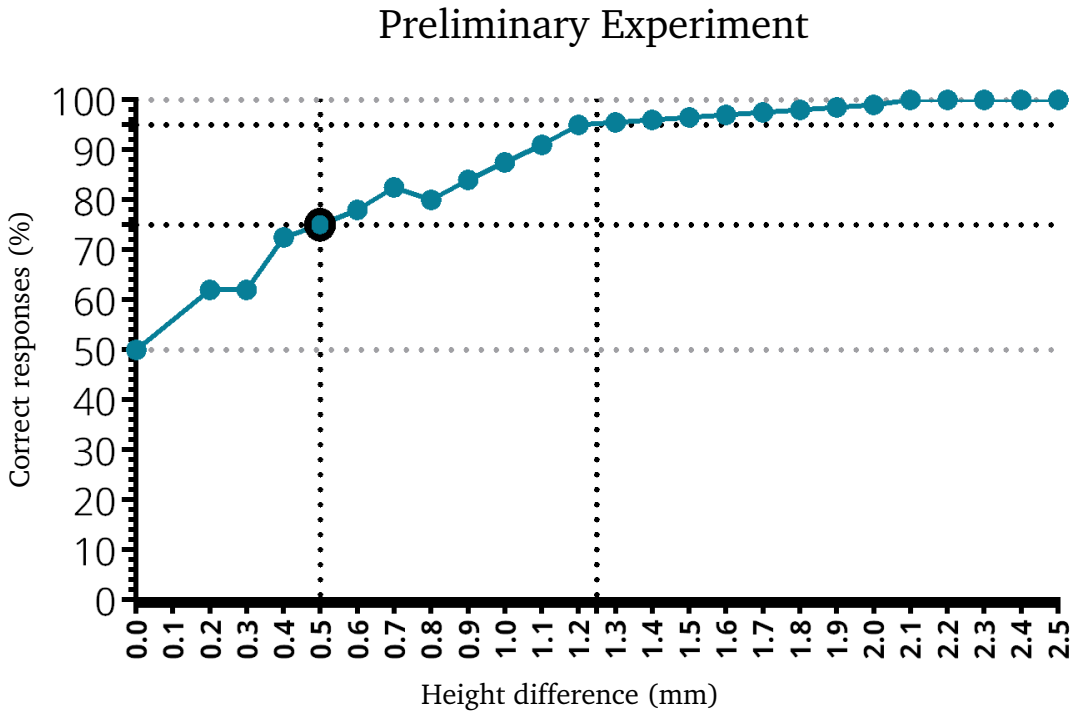


Figure 2.15: Correct Response Rate vs Height Difference

Comparison Experiment

The overall quantitative result from the series of experiments in the final phase is shown in Table 2.2 and illustrated in Figure 2.16. As seen in the psychometric graph, in

the presence of stereopsis, subjects' performance reached a clinically acceptable range ($> 90\%$[4]) when the height difference is greater than $0.75mm$. The same pool of subjects did not achieve the same threshold value for the height difference of less than $2.5mm$. The test of significance showed that the difference between the level of performance in stereo versus mono was indeed significant (paired t-test: $t = 2.93$, $p = 0.01$). In addition, all of our subjects preferred the stereo over the monocular system.

	Height difference (mm)										
	0	0.25	0.5	0.75	1.0	1.25	1.5	1.75	2.0	2.5	$\mu \pm \delta$
Mono	45%	60%	60%	70%	75%	79%	85%	85%	87%	90%	73.6 ± 14.6
Stereo	45%	45%	65%	90%	95%	98.2%	100%	100%	100%	100%	83.8 ± 23.0

Table 2.2: Average correct response rate (%) with respect to height difference (mm)

2.3.2 Phase 1: Discussion

Our data show that if the basilar artery is impinging on the membrane, deforming it by at least $0.75mm$, the stereopsis cue can allow the surgeon to avoid that area with 90% confidence. In addition to stereopsis, other monocular cues may be present, and consequently can be used to further improve the accuracy. However, the rate of performance in making use of monocular and stereo cues was shown to be subject-dependent. In other words, subjects have the ability to make use of one cue or the other, according to personal choice or perceptual capacity. In our experiment, however, the subjects were never worse when using stereo and mono cues, as compared with monocular vision alone; and in several cases, their acuity thresholds were improved significantly in the presence of stereopsis. Overall, stereoendoscopy was shown to facilitate the detection of critical structures (in the context of hydrocephalus patients), especially when the anatomical cues are poor.

[4] 90% threshold was chosen as clinically acceptable by our collaborator expert neurosurgeon.

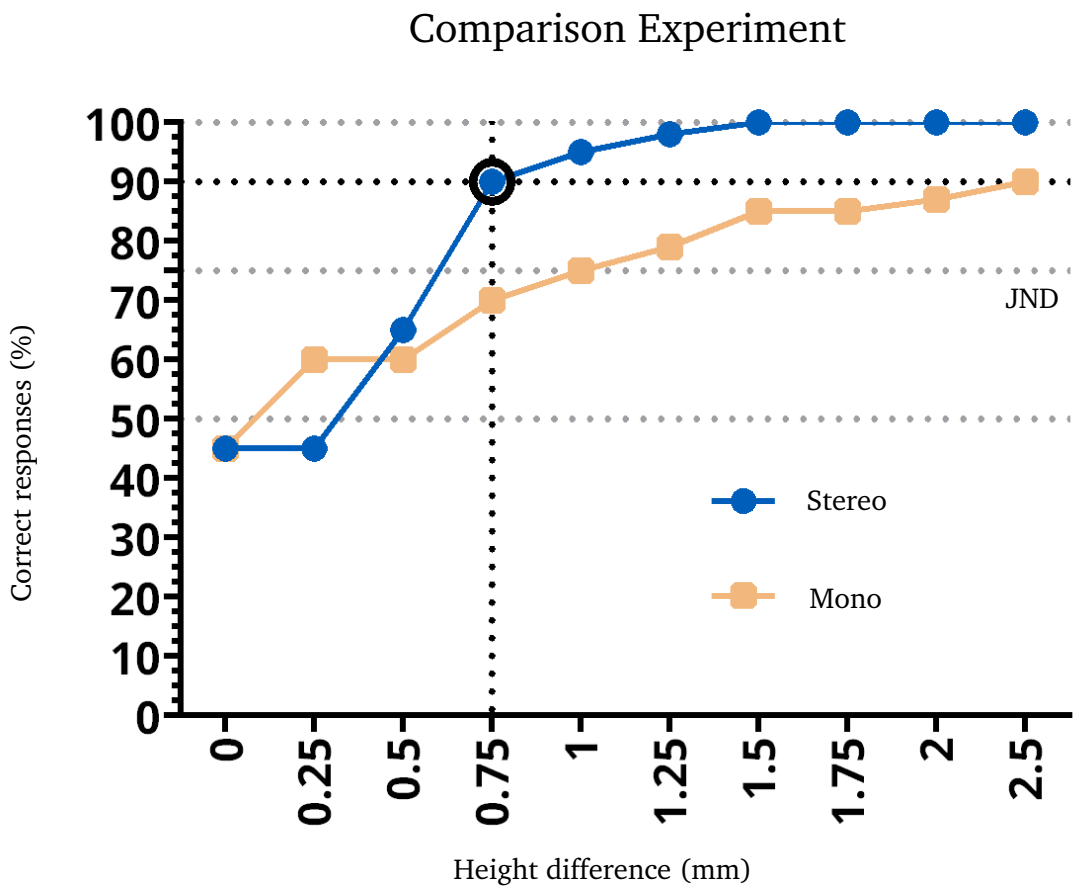


Figure 2.16: Correct Response Rate vs Height Difference

2.3.3 Phase 2: Results

Since the elevation above the basilar artery and clivus were offset randomly on the screen to avoid predicting the target location, each position was normalized by subtracting the actual location from the offset amount for further analysis. As a result, a histogram was produced across all trials, showing a distribution of target locations on a normalized scale (Figure 2.17). These data illustrate a larger variance of target distribution for novices. Despite such a large variance, however, novices performed within a clinically acceptable range ($\mu = 97.6\%$, $\delta = 2.08\%$) as shown in Table 2.3.

Subject	1	2	3	4	5	6	7	8	9	10	11	$\mu \pm \delta$
Expertise	high	low										
Hitting target in a safe zone (%)	100	100	100	100	98.6	98	97.3	96	96	96	94	97.8 ± 2.1

Table 2.3: Targeting performance (%) of novices/expert with the help of stereoscopic view

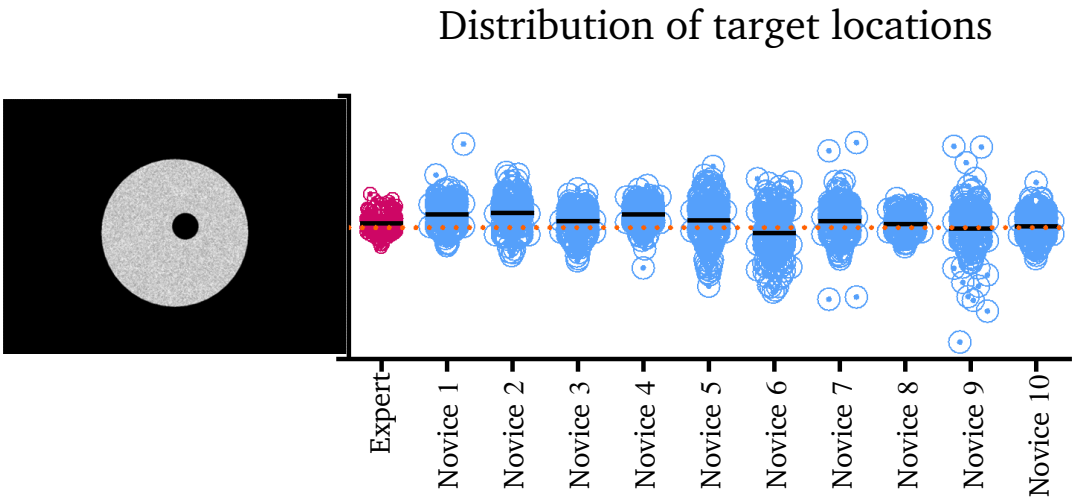


Figure 2.17: Novices show a larger variance of their targeting distributions compared to an expert

2.3.4 Phase 2: Discussion

Our results from the previous phase showed the stereopsis cue can be an essential asset to discriminate between two similar structures with different depth values. In phase 2, we implemented a paradigm whereby the location of stoma is selected based on visual cues that are off-target with an undefined visual zone. Our results show that the performance rate can be acceptably high (μ = 97.8%, δ = 2.11%) when individuals are only provided with the stereopsis cue, and that the expert tends to be more accurate in localizing a point as they demonstrate a significantly smaller standard deviation around their chosen mean target location.

2.3.5 General Discussion

There is no scientific consensus on the advantages of stereopsis in performing endoscopic procedures, and surgeons seem to perform endoscopic interventions remarkably well in the absence of stereopsis. This, however, can be a result of multiple factors, influencing the outcome. For instance, it is shown that depth perception is affected more by the surgeons' experience and anatomical knowledge rather than the method of visualization [39, 40], and that with enough training, trainees can form a mental image of anatomical structures and surgical instruments within the *displayed* environment, whether 2D or 3D. Therefore, it is not surprising that many studies have shown novice trainees may benefit more from stereopsis than experienced surgeons who have already developed mental and physical skills to operate under the 2D vision. In contrast, operating under the pseudo-3D vision provided by stereoendoscopes may initially cause expert surgeons visual and mental fatigue, degrading their performance. Thus, in order to make a fair comparison between mono- and stereoendoscopes, subjects, whether experts or novices, should undergo a similar training for both modalities prior to participating in experiments.

The second notion is the *task* performed. For most people, the visuomotor system relies heavily on binocular cues, as opposed to monocular ones, to perform controlled skilled reaching tasks[5] [41]. In contrast, tasks that involve judging the relative size depends upon our perception of the scene with both monocular and binocular cues affecting the outcome [42]. This phenomenon can be explained by the differentiation between perception and action observed in dorsal and ventral streams (refer to section 1.1.1 for more information). Furthermore, depending on the task, the operator's hand movements may appear rotated, mirrored, and amplified on a monitor that is placed in the distant location, all of which would affect the performance [10].

Thirdly, depending on the clinical application and environmental variables, stere-

[5]With an exception of people with ocular dominance

opsis is only useful when existing depth cues are either insufficient or may lead to visual misperception. When comparing monocular and stereoendoscopes, the optical and environment properties need to be adjusted such that they mimic the actual scenario, assuring the validity of the results. For example, improper lighting conditions may introduce shadows that can be perceived by the subject to estimate depth, underestimating the need for stereopsis.

Finally, stereoendoscopes can be combined with other approaches, tailored to the patient and pathology [43] to maximize their safety. For the most part, the clinical role of stereopsis has been overlooked due to successful adaptation of monocular endoscopes and technical limitations. We believe that with sufficient training and advances in technology, stereopsis can be an asset to surgical endoscopy in cases where other visual cues are impoverished and the surgical task can benefit from the 3D view. One example of such is the robotic-assisted laparoscopy procedures in which the surgical field is viewed under a 3D vision created by two optical channels. In the case of da Vinci® robot for example, stereopsis is shown to significantly improve the accuracy and reduce the operation time [44, 45, 46, 47].

2.4 Conclusion

The purpose of this study was to discover whether stereoendoscopy could improve neurosurgical targeting in third ventriculostomy. The nature of ETV operations requires surgeons to make the stoma precisely within a small area between the basilar artery and clivus. To perform this critical task, accurate localization of the basilar artery and clivus is necessary. We found that monocular endoscopes, although currently being used in ETV interventions, provide inadequate information for precise targeting when the floor of the third ventricle is thick and opaque. Stereoendoscopes, on the other hand, can be very helpful since they allow surgeons to locate the target area based

on the available depth information. Overall, this study demonstrates the capability of stereoendoscopes in terms of improving the neuro-surgical targeting and so increasing the patients' safety in third ventriculostomy.

2.5 Limitations and Future Work

Surgical targeting involves both *perception* and *action* with continuous feedback and feedforward mechanisms across them. Therefore, future work may involve conducting experiments within a realistic environment that involves (i) displaying surgical instruments in order to investigate the role of perception (e.g., stereopsis) on action (e.g., physically targeting a region-of-interest) and (ii) ensuring that 3D glasses are worn for a reasonable duration in order to examine the effect of visual fatigue on performance. Nevertheless, conducting experiments within a laboratory environment with the use of stationary stimuli and relatively low sample size prevent us from extending our results beyond the scope of the proposed model.

Bibliography

[1] National Institute of Neurological Disorders and Stroke. Hydrocephalus fact sheet, 2014.

[2] Triantafyllos Bouras and Spyros Sgouros. Complications of endoscopic third ventriculostomy: A review. *Journal of Neurosurgery: Pediatrics*, 7(6):643–649, 2011.

[3] Michelangelo Gangemi, Francesco Maiuri, Simona Buonamassa, Giuseppe Colella, and Enrico de Divitiis. Endoscopic third ventriculostomy in idiopathic normal pressure hydrocephalus. *Neurosurgery*, 55(1):129–134, 2004.

[4] P Grunert. From the idea to its realization: The evolution of minimally invasive techniques in neurosurgery. *Minimally invasive surgery*, 2013, 2013.

[5] P Cappabianca, A Alfieri, and E De Divitiis. Endoscopic endonasal transsphenoidal approach to the sella: Towards functional endoscopic pituitary surgery (feps)*. *min-Minimally Invasive Neurosurgery*, 41(02):66–73, 1998.

[6] Kai-Uwe Lewandrowski, Sang-Ho Lee, and Menno Iprenburg. *Endoscopic Spinal Surgery*. JP Medical Ltd, 2013.

[7] Dieter Hellwig, Wuttipong Tirakotai, Thomas Riegel, Stefan Heinze, and Helmut Bertalanffy. Endoscopy in neurosurgery. 104(4):185–91, 2007.

[8] Henry WS Schroeder. Current status and future developments of neuroendoscopically assisted neurosurgery. In *Neuroendoscopy*, pages 65–80. Springer, 2014.

[9] BL Bauer and D Hellwig. Minimally invasive endoscopic neurosurgerya survey. In *Minimally Invasive Neurosurgery II*, pages 1–12. Springer, 1994.

[10] P Breedveld, HG Stassen, DW Meijer, and LPS Stassen. Theoretical background and conceptual solution for depth perception and eye-hand coordination problems in laparoscopic surgery. *Minimally invasive therapy & allied technologies*, 8(4):227–234, 1999.

[11] Lawrence W Way, Lygia Stewart, Walter Gantert, Kingsway Liu, Crystine M Lee, Karen Whang, and John G Hunter. Causes and prevention of laparoscopic bile duct injuries: analysis of 252 cases from a human factors and cognitive psychology perspective. *Annals of surgery*, 237(4):460, 2003.

[12] Philip Servos, Melvyn A Goodale, and Lorna S Jakobson. The role of binocular vision in prehension: a kinematic analysis. *Vision research*, 32(8):1513–1521, 1992.

[13] Jörg Hofmeister, Timothy G Frank, Alfred Cuschieri, and Nicholas J Wade. Perceptual aspects of two-dimensional and stereoscopic display techniques in endoscopic surgery: review and current problems. *Surgical Innovation*, 8(1):12–24, 2001.

[14] MO Schurr, G Buess, W Kunert, E Flemming, H Hermeking, and L Gumb. Human sense of vision: a guide to future endoscopic imaging systems. *Minimally Invasive Therapy & Allied Technologies*, 5(5):410–418, 1996.

[15] JJ Marotta, Marlene Behrmann, and MA Goodale. The removal of binocular cues disrupts the calibration of grasping in patients with visual form agnosia. *Experimental Brain Research*, 116(1):113–121, 1997.

[16] Seth M Brown, Abtin Tabaee, Ameet Singh, Theodore H Schwartz, Vijay K Anand, et al. Three-dimensional endoscopic sinus surgery: feasibility and technical aspects. *Otolaryngology–Head and Neck Surgery*, 138(3):400–402, 2008.

[17] Rupali N Shah, W Derek Leight, Mihir R Patel, Joshua B Surowitz, Yu-Tung Wong, Stephen A Wheless, Anand V Germanwala, and Adam M Zanation. A controlled laboratory and clinical evaluation of a three-dimensional endoscope for endonasal sinus and skull base surgery. *American journal of rhinology & allergy*, 25(3):141–144, 2011.

[18] R Peter Manes, Sam Barnett, and Pete S Batra. Utility of novel 3-dimensional stereoscopic vision system for endoscopic sinonasal and skull-base surgery. In *International forum of allergy & rhinology*, volume 1, pages 191–197. Wiley Online Library, 2011.

[19] Garni Barkhoudarian, Alicia Del Carmen Becerra Romero, and Edward R Laws. Evaluation of the 3-dimensional endoscope in transsphenoidal surgery. *Neurosurgery*, 73:ons74–ons79, 2013.

[20] Abtin Tabaee, Vijay K Anand, Justin F Fraser, Seth M Brown, Ameet Singh, and Theodore H Schwartz. Three-dimensional endoscopic pituitary surgery. *Neurosurgery*, 64(5):ons288–ons295, 2009.

[21] Elina Kari, Nelson M Oyesiku, Vladimir Dadashev, and Sarah K Wise. Comparison of traditional 2-dimensional endoscopic pituitary surgery with new 3-dimensional endoscopic technology: intraoperative and early postoperative factors. In *International forum of allergy & rhinology*, volume 2, pages 2–8. Wiley Online Library, 2012.

[22] Paolo Castelnuovo, Paolo Battaglia, Mario Turri-Zanoni, Luca Volpi, Maurizio Bignami, and Iacopo Dallan. Transnasal skull base reconstruction using a 3-d endo-

scope: our first impressions. *Journal of neurological surgery. Part B, Skull base*, 73(2):85, 2012.

[23] Y-M Dion and F Gaillard. Visual integration of data and basic motor skills under laparoscopy. *Surgical endoscopy*, 11(10):995–1000, 1997.

[24] Klaus Peitgen, Martin V Walz, Markus V Walz, Gerald Holtmann, and Friedrich W Eigler. A prospective randomized experimental evaluation of three-dimensional imaging in laparoscopy. *Gastrointestinal endoscopy*, 44(3):262–267, 1996.

[25] N Taffinder, SGT Smith, J Huber, RCG Russell, and A Darzi. The effect of a second-generation 3d endoscope on the laparoscopic precision of novices and experienced surgeons. *Surgical endoscopy*, 13(11):1087–1092, 1999.

[26] C Von Pichler, K Radermacher, and G Rau. The state off 3-d technology and evaluation. *Minimally Invasive Therapy & Allied Technologies*, 5(5):419–426, 1996.

[27] Pirmin Storz, Gerhard F Buess, Wolfgang Kunert, and Andreas Kirschniak. 3d hd versus 2d hd: surgical task efficiency in standardised phantom tasks. *Surgical endoscopy*, 26(5):1454–1460, 2012.

[28] Konstantinos Votanopoulos, F Charles Brunicardi, John Thornby, and Charles F Bellows. Impact of three-dimensional vision in laparoscopic training. *World journal of surgery*, 32(1):110–118, 2008.

[29] Arjun Maini, Victoria Roach, Stephen Pautler, and Timothy Wilson. Can stereoscopic technologies build better surgeons? the effects of stereoscopy and spatial visualization ability on laparoscopic performance in surgical residents (536.12). *The FASEB Journal*, 28(1 Supplement):536–12, 2014.

[30] ACW Chan, SCS Chung, APC Yim, JYW Lau, EKW Ng, and AKC Li. Comparison of two-dimensional vs three-dimensional camera systems in laparoscopic surgery. *Surgical endoscopy*, 11(5):438–440, 1997.

[31] George B Hanna, Sami M Shimi, and Alfred Cuschieri. Randomised study of influence of two-dimensional versus three-dimensional imaging on performance of laparoscopic cholecystectomy. *The Lancet*, 351(9098):248–251, 1998.

[32] G Crosthwaite, T Chung, P Dunkley, S Shimi, and A Cuschieri. Comparison of direct vision and electronic two-and three-dimensional display systems on surgical task efficiency in endoscopic surgery. *British Journal of Surgery*, 82(6):849–851, 1995.

[33] Daniel B Jones, Jerome D Brewer, and Nathaniel J Soper. The influence of three-dimensional video systems on laparoscopic task performance. *Surgical Laparoscopy Endoscopy & Percutaneous Techniques*, 6(3):191–197, 1996.

[34] MD Mueller, C Camartin, E Dreher, and W Hänggi. Three-dimensional laparoscopy. *Surgical endoscopy*, 13(5):469–472, 1999.

[35] Elspeth M McDougall, Jon J Soble, J Stuart Wolf jr, Stephen Y Nakada, Osama M Elashry, and Ralph V Clayman. Comparison of three-dimensional and two-dimensional laparoscopic video systems. *Journal of endourology*, 10(4):371–374, 1996.

[36] Kyoung Won Nam, Jeongyun Park, In Young Kim, and Kwang Gi Kim. Application of stereo-imaging technology to medical field. *Healthcare informatics research*, 18(3):158–163, 2012.

[37] Hani J Marcus, Archie Hughes-Hallett, Thomas P Cundy, Aimee Di Marco, Philip Pratt, Dipankar Nandi, Ara Darzi, and Guang-Zhong Yang. Comparative ef-

fectiveness of 3-dimensional vs 2-dimensional and high-definition vs standard-definition neuroendoscopy: A preclinical randomized crossover study. *Neurosurgery*, 74(4):375–381, 2014.

[38] Tom N Cornsweet. The staircase-method in psychophysics. *The American journal of psychology*, pages 485–491, 1962.

[39] RS Sidhu, D Tompa, R Jang, ED Grober, KW Johnston, RK Reznick, and SJ Hamstra. Interpretation of three-dimensional structure from two-dimensional endovascular images: Implications for educators in vascular surgery. *Journal of vascular surgery*, 39(6):1305–1311, 2004.

[40] Kyle R Wanzel, Stanley J Hamstra, Marco F Caminiti, Dimitri J Anastakis, Ethan D Grober, and Richard K Reznick. Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. *Surgery*, 134(5):750–757, 2003.

[41] Philip Servos. Distance estimation in the visual and visuomotor systems. *Experimental Brain Research*, 130(1):35–47, 2000.

[42] David C Knill. Reaching for visual cues to depth: The brain combines depth cues differently for motor control and perception. *Journal of Vision*, 5(2):2, 2005.

[43] Jonathan Roth, Ameet Singh, Gurston Nyquist, Justin F Fraser, Antonio Bernardo, Vijay K Anand, and Theodore H Schwartz. Three-dimensional and 2-dimensional endoscopic exposure of midline cranial base targets using expanded endonasal and transcranial approaches. *Neurosurgery*, 65(6):1116–1130, 2009.

[44] Gyung Tak Sung and Inderbir S Gill. Robotic laparoscopic surgery: a comparison of the da vinci and zeus systems. *Urology*, 58(6):893–898, 2001.

[45] Darius Burschka, Jason J Corso, Maneesh Dewan, William Lau, Ming Li, Henry Lin, Panadda Marayong, Nicholas Ramey, Gregory D Hager, Brian Hoffman, et al. Navigating inner space: 3-d assistance for minimally invasive surgery. *Robotics and Autonomous Systems*, 52(1):5–26, 2005.

[46] Adélaïde Blavier, Quentin Gaudissart, Guy-Bernard Cadière, and Anne-Sophie Nyssen. Impact of 2d and 3d vision on performance of novice subjects using da vinci robotic system. *Acta Chirurgica Belgica*, 106(6), 2006.

[47] Y Munz, K Moorthy, A Dosis, JD Hernandez, S Bann, F Bello, S Martin, A Darzi, and T Rockall. The benfits of stereoscopic vision in robotic-assisted performance on bench models. *Surgical Endoscopy And Other Interventional Techniques*, 18(4):611–616, 2004.

Chapter 3

Visual Enhancement of MRA Images to Facilitate Planning of AVM Interventions

This chapter is adapted from '*Visual Enhancement of MR Angiography Images to Facilitate Planning of Arteriovenous Malformation Interventions*'[1].

My contribution to this chapter involved (i) designing and conducting experiments, (ii) analyzing data, and (iii) writing manuscripts.

3.1 Introduction

3.1.1 Clinical Motivation: Arteriovenous Malformations

Arteriovenous malformations (AVMs) are vascular lesions in which blood flows directly from the arterial to the venous system bypassing the normally interposed capillary bed.

[1]Abhari, K., Baxter, J., Khan, A., Peters, T., de Ribaupierre, S., Eagleson, R. "Visual Enhancement of MR Angiography Images to Facilitate Planning of Arteriovenous Malformation Interventions", Revised and Resubmitted, Submission No. TAP-2014-0021.R1, *ACM Transactions on Applied Perception*, (2014)

It is estimated that 300,000 Americans are affected by AVMs [1]. Symptoms of AVMs include seizures and headaches, and have an associated 2-4% risk of hemorrhage per year [2]. Treatment is multimodal involving a combination of one or more of the following interventions: endovascular embolization, stereotactic radiation, and open surgery [3] [4]. If surgical resection is chosen, preoperative planning is necessary to localize the anomaly and understand relevant neurovascular anatomy through examination of the patient's angiography images. These images are acquired preoperatively by means of magnetic resonance angiography (MRA), computed tomography angiography (CTA), or cerebral angiography. Regardless of the method of acquisition, blood vessels are visually distinct in angiographic images due to relatively high contrast between the vascular lumen and surrounding tissues.

The aim of surgical planning is to identify the optimal path and point of entry, as well as to localize and differentiate the arteries from the veins in order to avoid intraoperative hemorrhage. Planning can be very challenging because 30%-70% of AVM patients harbor this anomaly in functional areas of their brain [5] limiting the number of possible surgical paths and entry points. In addition, AVMs are vascular abnormalities in which the arteries and veins are twisted together making differentiation and localization of vessels extremely difficult (Figure 3.1). Each case of AVM is also unique making previous anatomical knowledge unreliable. Nevertheless, planning AVM resection interventions, if performed by conventional means, can be a cognitively challenging task because of the following shortcomings:

1. In the current standard of care, CT or MR angiography images are visualized either in a (2D) slice-by-slice or 3D volumetric fashion. In the former, interaction involves scrolling through a sequence of 2D slices of preoperative scans presented on the axial, sagittal, and coronal planes (Figure 3.2). This method of interaction can require significant cognitive effort, that is the engaged proportion of the working memory, to form a mental 3D representation of the brain, the lesion,

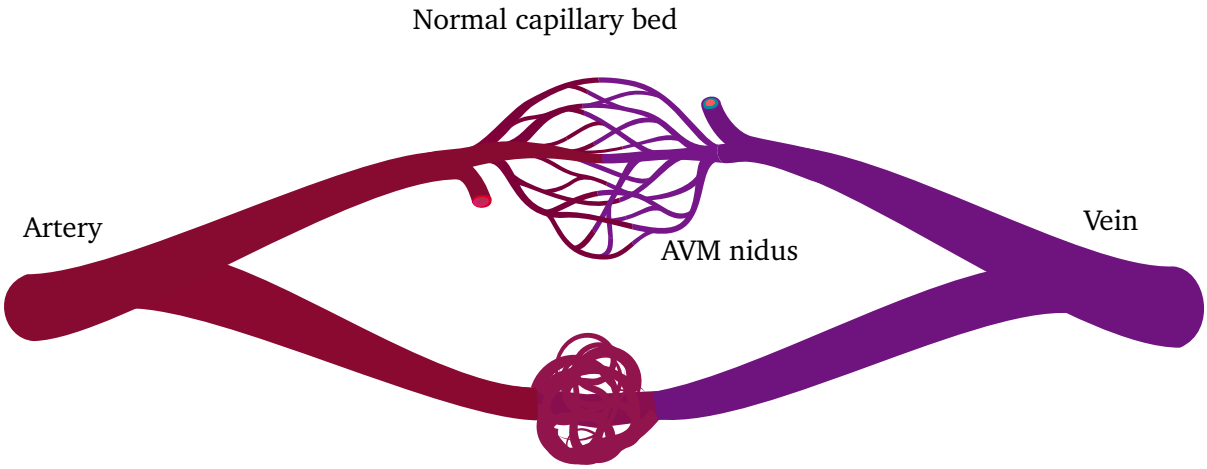


Figure 3.1: AVM manifests itself as a tangled mass of blood vessels

and surrounding eloquent structures, and to locate, identify, and follow certain vessels from one slice to another. Furthermore, these methods do not allow for the visualization of different brain structures within the same plane or exploring them from an arbitrary point of view, making a proper risk assessment difficult. Even though interacting with 3D volumetric images is more intuitive, the method of rendering is a determining factor in successful planning (this is discussed in details in section 3.2.1).

2. The center of AVM malformations - commonly referred to as the *nidus* - usually manifests itself as a cluttered mass of blood vessels (Figure 3.1), which can increase the time and error rate for judging the relative depth and spatial relationship between the vascular structures [6].

With the goal of facilitating AVM preoperative planning, we introduce a number of computer graphics techniques to address the aforementioned perceptual issues, and then discuss how we can evaluate these techniques using pyschophysical tests.

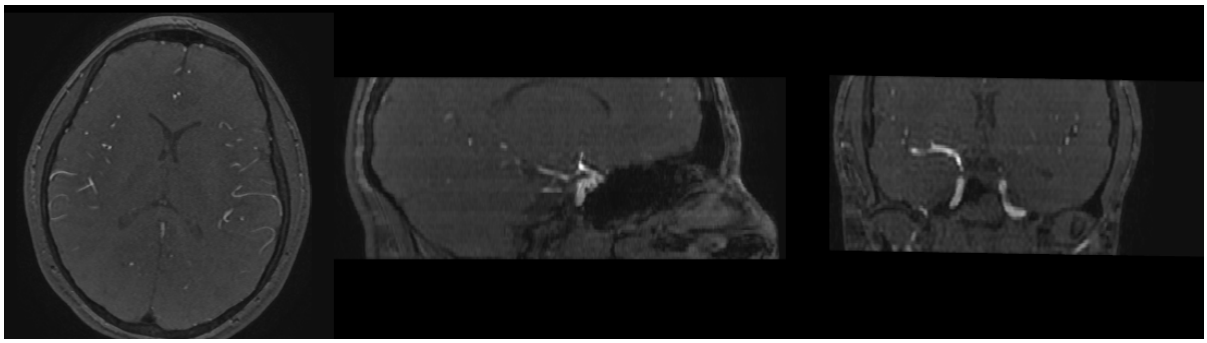


Figure 3.2: Conventional representation of MRA images on axial, sagittal, and coronal planes

3.1.2 Background

Related Work

The manner by which vascular structures are visualized in angiographic images is a critical factor in the success of an AVM intervention, and is thus an emerging area of research. Bullitt et al., for instance, developed an application to visualize the angio-architecture of AVM lesions [7]. In their study, vascular trees are segmented by placing seed points to initiate extraction, and then colour-coded by incorporating information about their parent-child relationships. Weiler et al. have proposed a processing pipeline whereby feeding arteries and draining veins are detected and colour-coded by acquiring and processing a combination of MR images [8]. Chen et al. made use of computed tomography digital subtraction angiography and 3D X-ray angiography to detect and label arteries and veins using level-set flow propagation, and then decluttering the nidus using the topological distance between vessels [9]. These studies, although beneficial, may not be always practical or cost-effective for routine clinical settings. For example, robust vascular segmentation may require extensive time and user effort. Relying on certain imaging modalities - such as MR Venography - to detect and extract structures could be inaccurate in that it may ambiguously or incorrectly label a vessel

[9]. Furthermore, non contrast-enhanced MRA images are often preferred over modalities that involve ionizing radiation (e.g. 3D X-ray and CTA) or injection of contrast media (e.g. CE-MRA and cerebral angiography). For these reasons, computer graphics techniques have been proposed to improve the perception of relative depth among vascular branches. Ropinski et al. enhanced depth perception in cerebral angiography images by introducing colour, depth of field, and contours [10] . They also explored the effect of enhancing contours with the thickness varying inversely to the distance from the viewer. Their study indicated that the overall performance (accuracy, time, and subjective feedback) was improved when colour and depth of field were used. The lowest user error, however, was achieved when edges were highlighted with constant thickness value. Joshi et al. made use of halos, distance colour blending, and tone shading to enhance the perception of depth in volumetric images of vascular structures [11]. According to their evaluation study, distance colour blending was preferred by experts over other techniques. In [12], a number of non-photorealistic algorithms such as hatching and distance-encoded shadows were used to encode depth information in isosurface vascular structures. Their proposed technique has been embedded into an augmented reality environment for use in liver surgery [13]. Their method, however, is only applicable to segmented structures. Kersten-Oertel et al. performed an extensive evaluation study in which the effect of certain visual cues (stereo, local occlusion via depiction of edges, aerial perspective, kinetic depth, and pseudo-Chromadepth[2]) on establishing the relative depth between different vessels was explored [14]. According to their results, both novices and experts performed significantly better when the pseudo-chromadepth and fog cues were presented. Edge enhancement resulted in improved performance in case of experts only. In contrast, stereoscopic and kinetic depth cues did not yield good results, neither in time nor in correctness. In this work, human studies are conducted to investigate the role of contours and stereopsis, and their possible

[2]Chromadepth is a stereoscopic effect generated by prismatic filters separating the image into the visible spectrum. Pseudo-Chromadepth is Chromadepth with limited visible light spectrum [14].

interaction, in the perception of *connectivity* and *relative depth* of vascular structures in *non contrast-enhanced MRA* (NC-MRA, Time-of-Flight imagery) images.

Perceptual Factors

Contours: Early stages of visual perception involve unconscious, rapid accumulation of information (200 msec - 250 msec [15]) provided by *pre-attentive features*. These features are detected in parallel [16] and used by our brain to direct attention and to focus top-down visual processes [17]. Pre-attentive features are often classified into four categories: colour, movement, spatial localization, and *form* [18]. Recognizing forms - that include dots, lines, shapes, and *contours* - is the basis of object recognition. Contours, in particular, help us recognize shape, orientation, and relative depth, especially when other visual cues are impoverished. Locations in the image where one contour occludes another are evident as line terminators and edge-based features, providing powerful cues to relative depth. This information is then integrated with other information concerning motion, colour, and depth which is then matched with structural descriptions in memory.

Recognizing objects within a visually cluttered scene is a difficult task. Clutter in 3D leads to visual occlusion which degrades perceptual performance [19], particularly when little or no visual contrast exists between objects. One such example is identifying vessels-of-interest within a cluttered nidus. Nevertheless, similar studies have shown that if attention is directed to certain parts of an object, particularly when it is partially occluded, other parts of the same object also benefit from such attentional directedness [20]. In the case of AVMs, contour enhancement may therefore not only accentuate pre-attentive perception, but also direct the attention to certain parts of the nidus, facilitating top-down perception of the nidus itself.

While many medical image visualization studies have explored the use of computer

graphics techniques to improve the realism of tissue representation [21][22][23][14], the effect of non-photorealistic techniques to enhance visual features such as contours in the quality of diagnosis or intervention is still an open question.

Stereopsis: Depth perception involves consolidation of different monocular and binocular cues to form a three-dimensional representation of the world from the two-dimensional images projected onto the retina. Binocular cues such as *stereopsis* provide information about depth when a scene is observed with both eyes. Stereopsis is the process of inferring depth from two adjacent views of the world projected onto the left and right retinas with horizontal difference between these two retinal images, known as *binocular disparity*, being employed by disparity-tuned cells in several cortical visual areas to infer depth. Stereopsis is an important visual cue for depth perception, but only over short distances [24]. *Stereoacuity* is the sensitivity of stereopsis in representing small disparity differences. Luminance, spatial frequency, observation time, and *contours* among others can influence stereoacuity.

Despite the fact that stereopsis has been studied extensively in the literature (e.g. [25], [26]), the effect of contours and stereopsis both individually and in combination on the perception of medical images has not yet received enough attention from the medical image visualization community.

3.1.3 Hypothesis and Objectives

We believe that planning AVM interventions can be significantly improved by rendering vascular structures in 3D while enhancing their surrounding contours. We hypothesize that enhancing contours in particular enables users to trace the encompassed vascular structures with higher accuracy and efficiency, and capture relative spatial relationship between different vascular structures and between vascular structures and adjacent anatomical regions.

The objective of this study is to explore the role of contours and stereopsis in perceiving the *connectivity* and *relative depth* of vascular structures in *non contrast-enhanced MRA* (NC-MRA - Time-of-Flight imagery). We deliberately chose NC-MRA images due its safety, wider availability, and lower vascular visibility compared to contrast-enhanced MRA, cerebral angiography, or CTA images. NC-MRA is often used when patient safety is concerned as it is non-invasive and does not involve ionizing radiation or injection of contrast media.

In this study, a combination of volume rendering (section 3.2.1) and a non photo-realistic shading model (section 3.2.1) is proposed to visualize and enhance vascular structures. Quantitative (sections 3.3.1 & 3.3.1) and qualitative experiments are conducted to explore and examine our hypothesis.

3.1.4 Contributions

Our main contributions can be summarized as follows:

1. Non-photorealistic shading techniques are often developed for isosurfaces. In this study we extended a non-photorealistic algorithm, commonly known as cel-shading 3.2.1, to enhance the contours of *volumetric* images routinely acquired in medical applications;

2. Most related studies involve imaging modalities that provide very good contrast often at the cost of patient safety due to radiation. In this study, we made use of non contrast-enhanced MRA images with relatively low image contrast, and yet were able to show the significance of perceptual enhancement on performance; and

3. Most related studies have investigated the effect of contours on depth perception. In this work, we also explored the role of contours in the perception of continuity of vessels in clinical images.

3.2 Materials and Methods

3.2.1 Visualization of Vascular Structures

Direct volume rendering vs. Maximum Intensity Projection

Volume rendering involves the rendering of serially acquired 2-dimensional slices of medical datasets in a 3-dimensional space. One simple form of volume rendering, *Maximum Intensity Projection* (MIP), is commonly used for visualizing blood vessels in angiographic images. MIP has been employed for more than two decades taking advantage of its similarity to X-ray angiography, simplicity, and fast processing speed. In MIP, virtual rays are cast from the camera's viewpoint, but only the voxels with the highest intensity along rays are projected onto the 2D viewing plane [27], which in turn, results in a set of 2D images where blood vessels are visibly represented. One immediate drawback of MIP is that the 3D relationships between vessels are not preserved due to the rendering process and lack of shading [28]. For instance, in MIP images where veins and arteries are both depicted, the location of veins may appear more posterior than their actual anatomical location because of their lower intensity [28].

Unlike MIP, *Direct Volume Rendering* (DVR) using *ray casting* relies on the efficient computation of a compositing integral for each pixel in the output image [29] [30]. In this technique - similar to MIP - rays are cast outward from the viewing plane into the virtual space. The rendering equation combines colour and opacity information along a series of sample points spaced along the rays. Information about the colour/opacity is stored in a *transfer function* (TF), indexed by local information. The most common transfer functions are one-dimensional and use only intensity information from the image as the index. Two-dimensional transfer functions (2D-TF) employ both intensity and its gradient magnitude as a pair of indices for more complex rendering. The application of 2D-TFs in reducing the ambiguity and conveying subtle surface properties has been illustrated in the literature [21]. The result of the rendering equation is the

projection of the 3D volume on the screen.

Using DVR to render angiography images, clinicians can visualize blood vessels individually, prevent misdiagnosis [28], preserve the spatial relationship between different anatomies and structures, observe overlaying structures within the same visual plane, and avoid the time-consuming segmentation needed for surface rendering. Despite this, DVR is not yet recognized as a standard visualization approach for many medical applications. This could be because adjusting TFs can be cumbersome with results varying considerably from user to user, and processing large volumetric datasets can be computationally costly. Nevertheless, improvements in GPU programming have made it possible to render large volumes in real-time, and the preference of MIP over DVR is becoming questionable, particularly in neuroclinical interventions where understanding the relative relationship among intracranial vessels (e.g. AVM, Vein of Galen malformation), or between vessels and other brain structures (e.g. vascularized tumours) is crucial.

Cel-shading

Cel-shading (also known as *toon-shading*) is a non-photorealistic technique that involves highlighting contour lines with darker segments to emphasize edges. In this technique, rays are cast from the camera's viewpoint to sample data points within the volume of interest. Among these rays, the depth (or Z-value) of the last salient voxels are recorded. An *edge* is then defined where the difference in depth between a data point and its neighbors is greater than a certain threshold set by the user. In other words, the contours of volumetric objects are defined at discontinuities in the depth buffer (or the Z-buffer) as viewed by the camera (Figure 3.3). In this study, a sigmoid shape function is employed to control the highlighted discontinuity. In addition to delineating contours, increasing the distance between two objects - i.e. the relative depth between them - introduces more emphasis around the edge separating the two.

The use of DVR to visualize different tissue structures inside a *translucent* context, such as vascular structures inside the brain, often creates a foggy cloud around opaque objects disrupting the view. Since our algorithm for contour enhancement is not solely sensitive to ray penetration but rather to discontinuities of the Z-buffer, contour delineation provides a very strong cue to the connectivity and local structure of opaque models inside cloudy areas.

Gradient Shading

Throughout this chapter, the term cel-shading refers to the use of the Z-buffer based algorithm for enhancing contours, and so *cel-shading darkness* and *contour enhancement* are interchangeable. Cel-shading, in a broader context, involves a non-photorealistic illumination-reflection model that is not employed in this study. Instead, the illumination-reflection is computed as a function of normalized gradient vectors, and referred as *gradient-shading* throughout this paper. In this method, gradient vectors are used to approximate the surface normals in volumetric data. The overall intensity of each voxel is then calculated by a combination of a gradient model and the opacity value derived from the 2D-TFs. The relative computational efficiency of this approach makes the proposed visualization suitable for real-time applications.

An example of the gradient- and cel-shading models are shown in Figure 3.4.

3.3 Evaluation Studies

3.3.1 Experiment Design

Continuity: A Perceptual Study

The objective of this study was to investigate the role of contour enhancement on the

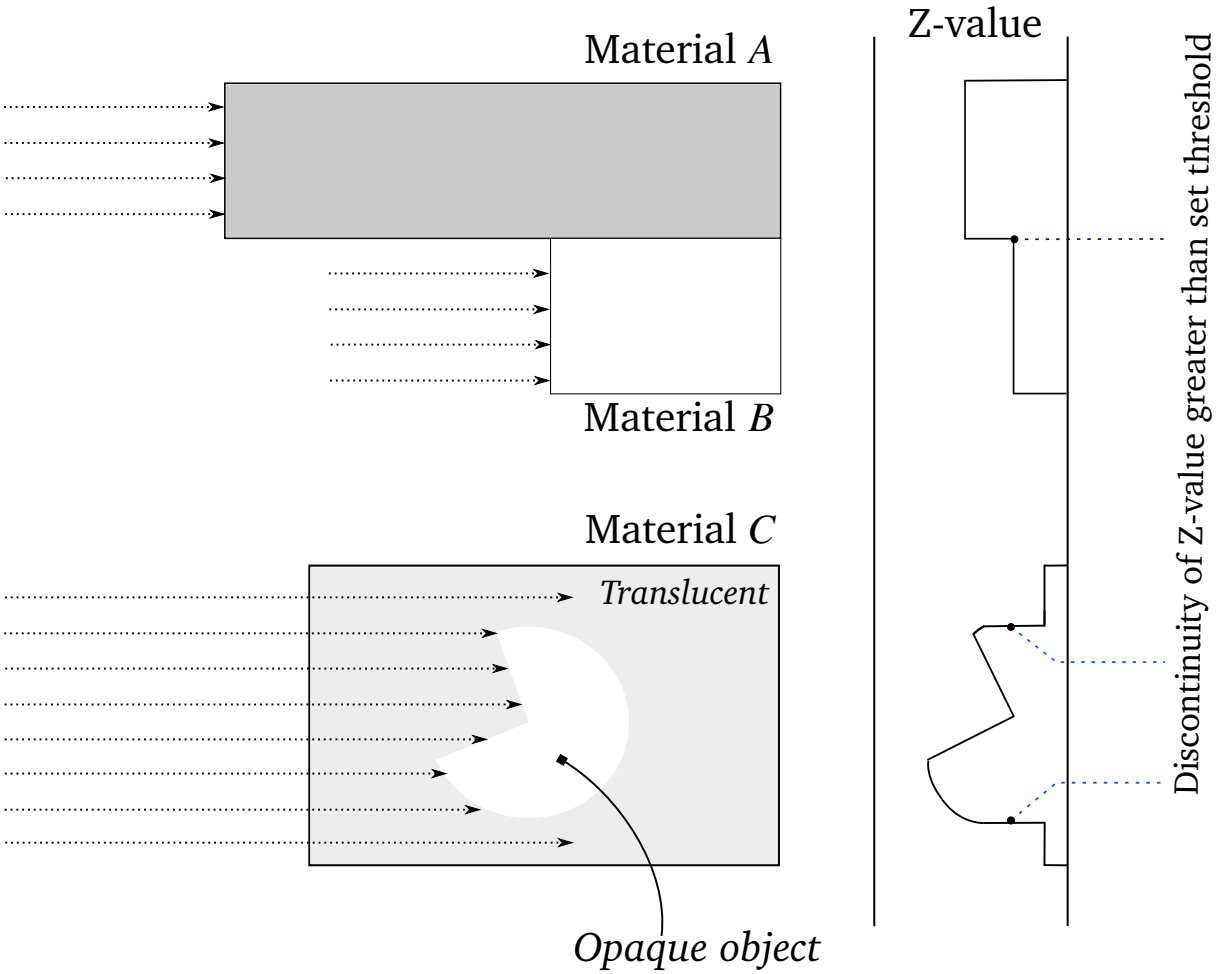


Figure 3.3: Top: Z-buffer edge enhancement is based on the discontinuities of the Z-value, relative to the viewer; Bottom: Opaque objects can be visualized if they are placed within translucent materials

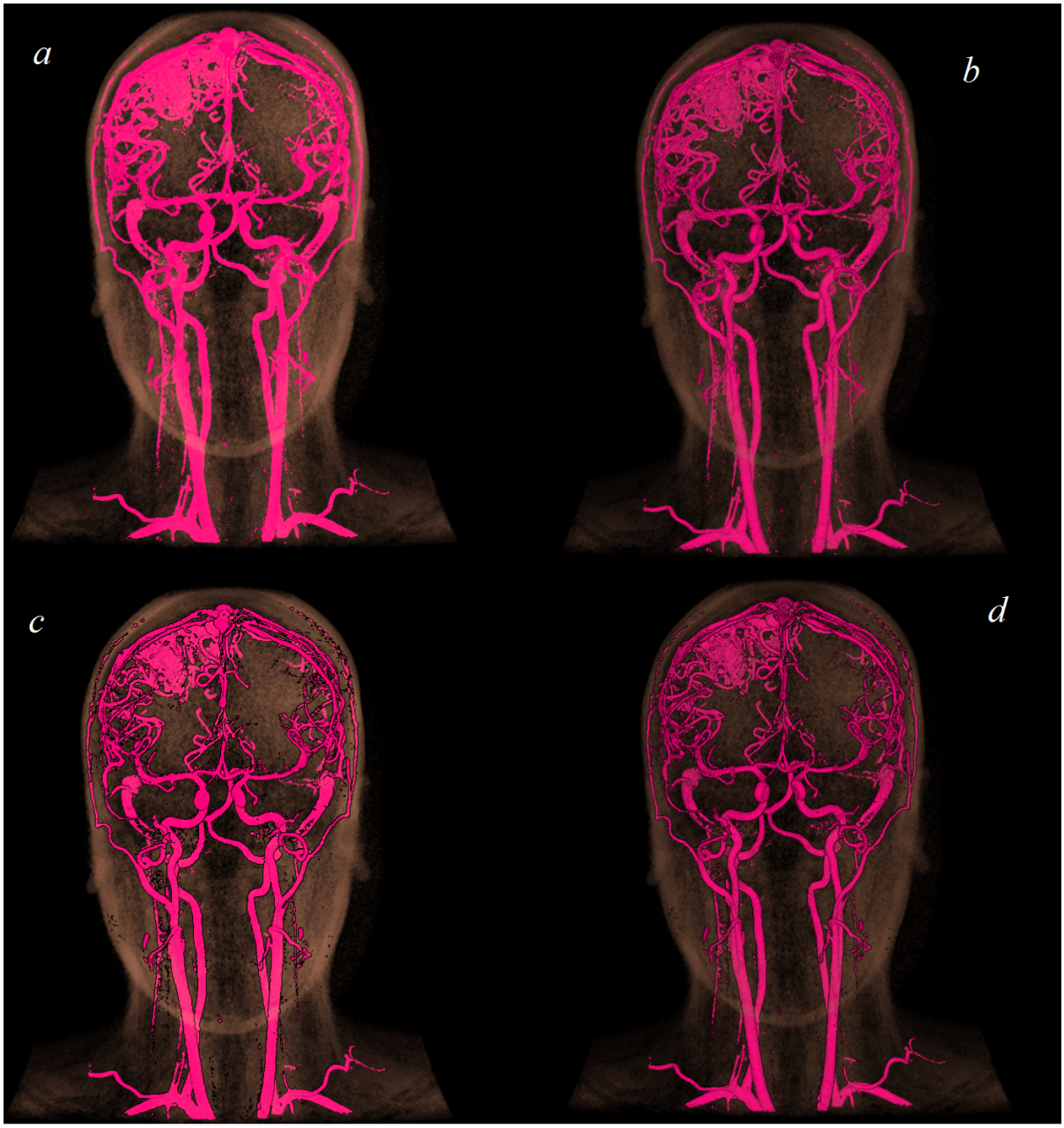


Figure 3.4: Different shading models: (*a*) no shading, (*b*) gradient-shading only, (*c*) cel-shading only, and (*d*) gradient and cel-shading combined.

perception of vessel continuity in NC-MRA images. Experiments involved 180 trials per subject in which subjects (3 female/7 male graduate students) were presented with a randomized series of stimuli, displayed on an LCD monitor[3]. Stimuli consisted of 30 stationary snapshots of volume-rendered NC-MRA images with and without cel-shading applied, generating 60 stimuli per subject. An expert neurosurgeon designed the 2D-TF and chose the angle and the field of view to reproduce the view typically seen in clinical practices. During the experiments, participants were asked to report whether or not two vessels (or two parts of one vessel) were connected, which were indicated by two colour-coded circles (Figure 3.5). Individuals were also asked to respond as quickly and accurately as possible while their responses and response times were recorded for further analysis. The images were systematically placed in sequence such that the effects of learning and fatigue did not provide an advantage or disadvantage to either case. The true connections between the target vessels were determined by exploring the 3D volume-rendered MRA images and confirmed by a neurosurgeon who examined the data by exploring the 2D NC-MRA scans. At the end of each experiment, subjects were also asked to provide us with their subjective assessments. Furthermore, each individual performed the test only once.

Depth: A Psychophysical Study

The objectives of this study were to determine:

1. Whether stereopsis facilitates the perception of relative depth in the context of blood vessels;

2. Whether contour enhancement facilitates the perception of connectivity and relative depth in the context of blood vessels;

[3]24" LG Flatron W2442PA with the resolution of 1920×1080

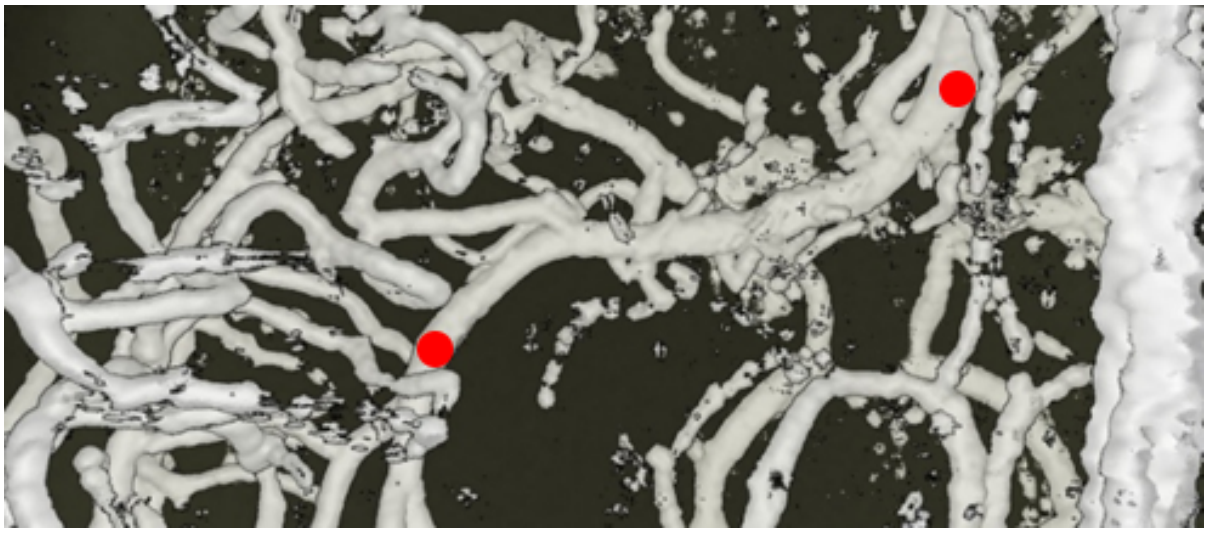


Figure 3.5: Circles indicate the vessel(s)-of-interest. In this example, edges are emphasized and target vessels are connected

3. The optimal combination of disparity and cel-shading salience to achieve improved performance, which can then be used in our cel-shading model to optimize for contour perception.

In our experiments, novice subjects (3 female/6 male graduate students) were asked to sit comfortably in front of a 3D display[4] and perform a task as quickly and accurately as possible. A training session was conducted at the beginning of each experiment to familiarize participants with the environment and their assigned task. The experiments consisted of a number of sequential trials (n= 336 per subject) in which a fixation cross was presented for 500 ms followed by a stimulus composed of two vertical pairs of cylindrical bars parallel to each other (Figure 3.7 - *left*). The fixation cross maintains the attention to the middle of screen avoiding eye movements that can potentially increase reaction time and/or confound the effects of attentive vision.

The experimental design involved a discrimination task based on the two-alternative forced-choice (2AFC) paradigm. In this paradigm, a pair of stimuli consisting of a target and a lure are presented in every trial. The observer's task is to report the stimulus

[4]54.6" Samsung 8000 series 3D LED TV with active shutter glasses and the resolution of 1920×1080; stereopsis with vertical interlaced was used, generating left/right images with the resolution 960×1080.

that contains the target. Unlike yes-no paradigms where subjects must report the presence or absence of a stimulus, forced-choice experiments are shown to be independent of the observers' subjective report of their personal perception [31]. Based on this paradigm, the task was defined as judging the relative depth between the two pairs of cylindrical bars (Figures 3.6 and 3.7) with the help of apparent visual cues, specifically stereopsis and occlusion. The target was defined as the pair with larger relative depth. Placing the bars at three different locations (near, middle, or far) relative to the camera created three different levels of disparity, and therefore stereopsis (distance from the cameras: 118, 342, and 566 cm respectively). This range of disparity was determined by an expert to be clinically relevant. In addition to stereopsis, *contour darkness*[5] was also alternated from one trial to the next generating contours with 3 different levels of enhancement: (i) *no enhancement*, (ii) *full enhancement* at which the amount of brightness of voxels around the edge was set to 0, making them appear fully dark, and (iii) *medium enhancement* at which the amount of brightness of voxels around the edge was set to 50% of their brightness value. Following such protocol resulted in 12 different conditions of different degrees of stereopsis and contour darkness (Table 3.1).

The presentation of these stimuli was varied to counterbalance within-subjects factors such as learning and fatigue. This was accomplished by dividing subjects into two groups; within each group, the experiment comprised six different blocks; and in each block, stimuli were presented with different amount of stereopsis and/or darkness:

1^{st} *group* : *darkness* $\rightharpoonup$ *stereo* $\rightharpoonup$ *stereo & darkness* $\rightharpoonup$ *stereo & darkness* $\rightharpoonup$ *stereo* $\rightharpoonup$ *darkness*

2^{nd} *group* : *stereo* $\rightharpoonup$ *darkness* $\rightharpoonup$ *stereo & darkness* $\rightharpoonup$ *stereo & darkness* $\rightharpoonup$ *darkness* $\rightharpoonup$ *stereo*

[5] In this chapter, darkness refers to the opposite of brightness, that is absence of light

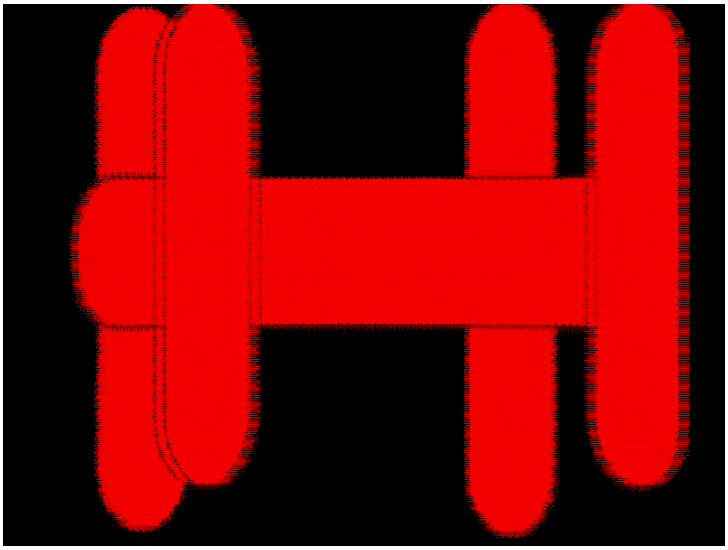


Figure 3.6: An example of stimuli with stereopsis (interlaced) enabled and contours enhanced.

Although the presentation of stimuli were randomized within each block, fully randomizing the sequence of blocks was avoided to mitigate for learning effect that might have occurred in trials where both visual cues existed. For analysis purposes, each trial was tagged according the relative depth between the cylinders.

In order to avoid vertical parallax and produce geometrically-correct views, images were rendered using perspective projection and stereo pairs were generated using off-axis method (as opposed to toe-in). However, perspective projections may result in perspective cue, biasing the results. To avoid this, half of the stimuli involved larger bars in the background, making *relative size* a misleading cue and so producing not more than 50% success if chosen randomly (Figure 3.8).

3.4 Results and Discussion

3.4.1 Continuity Experiment

Among the 10 subjects, 9 performed better when presented with contour enhanced images. The success rate, that is the overall score of correctly connecting the target

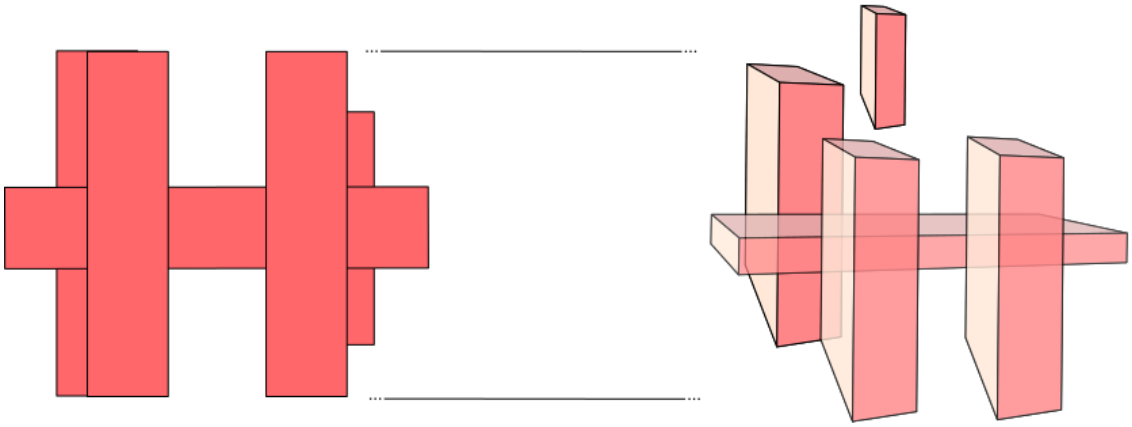


Figure 3.7: Left: Schematic presentation of stimuli (observers point of view); Right: The relative depth between near and far vertical bars (side view)

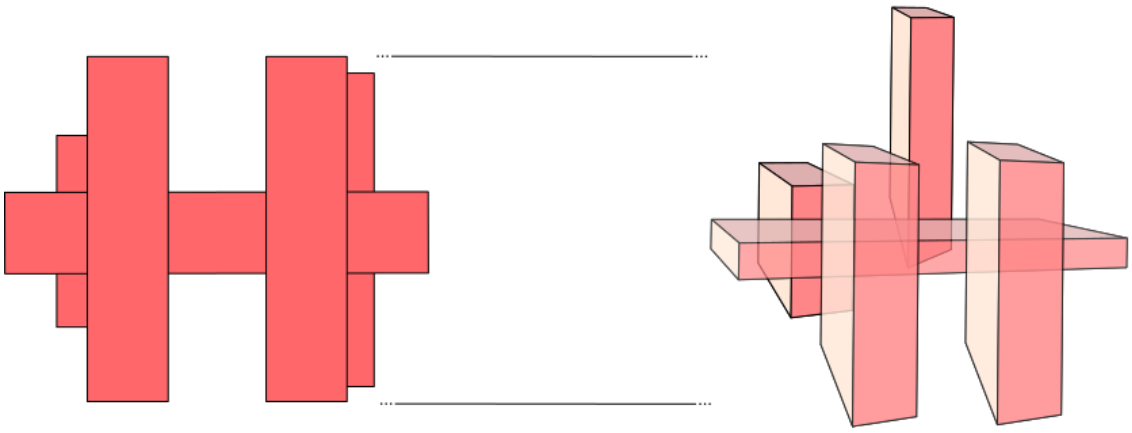


Figure 3.8: The bars in the background were generated larger in size compared to those in the foreground to avoid reliance on relative size as a visual cue

	Disparity (Distance to Camera)			
Darkness (Cel-shaded)	0 (*None*)	1.85^o (*Near*)	1.80^o (*Middle*)	1.69^o (*Far*)
None (0×max.)	1	2	3	4
Medium (.5×max.)	5	6	7	8
Full (1×max.)	9	10	11	12

Table 3.1: 12 different conditions of different degrees of stereoscopic disparity and cel-shaded darkness

points, was improved by 8.4% in contour-enhanced trials compared to non-enhanced trials ($\mu \pm \delta = 84\% \pm 34.4\%$ (enhanced) vs. $75.6\% \pm 41.7\%$ (non-enhanced); $p = 0.02$ (paired t-test), Cohen's d [32] $= 0.87$, Statistical power $= 0.89$, Figure 3.9). The overall response times (RT) were reduced significantly in enhanced trials ($\mu \pm \delta = 11.1 \pm 7.2$ sec (enhanced) vs. 13.4 ± 9.5 sec (non-enhanced); $p = 0.04$ (paired t-test), Figure 3.10). Conducting a Hotelling t-squared test showed the results were indeed significant considering both accuracy and RT as dependent measures of performance ($p = 0.01$, Hotelling t-squared $= 8.85$).

The level of difficulty in detecting the target from background events can be measured as the proportion of targets that are correctly identified as such (i.e. (*hits*) / (*hits*+*miss*)), that is the *sensitivity* of the system. Sensitivity is therefore the degree of separation between targets and background events, or as described in the *signal detection theory*, between *signal* and *noise*. According to the signal detection theory, the *sensitivity index* - commonly referred to as d' - is a measure of correlation between the level of task difficulty and the distance between signal and noise distributions. In experiments with *yes-no* paradigm, d' provides a standard measure of performance as: $d' = (Z_{hit_{rate}} - Z_{FA_{rate}})$, where $Z_P(p \in [0\ 1])$ is the inverse of the cumulative distribution function. Naturally, higher d' values denote improved perceptual performance in detecting signal among noise. In our experiments, the difficulty of locating and connecting target vessels can be estimated by comparing the sensitivity, or d', of enhanced

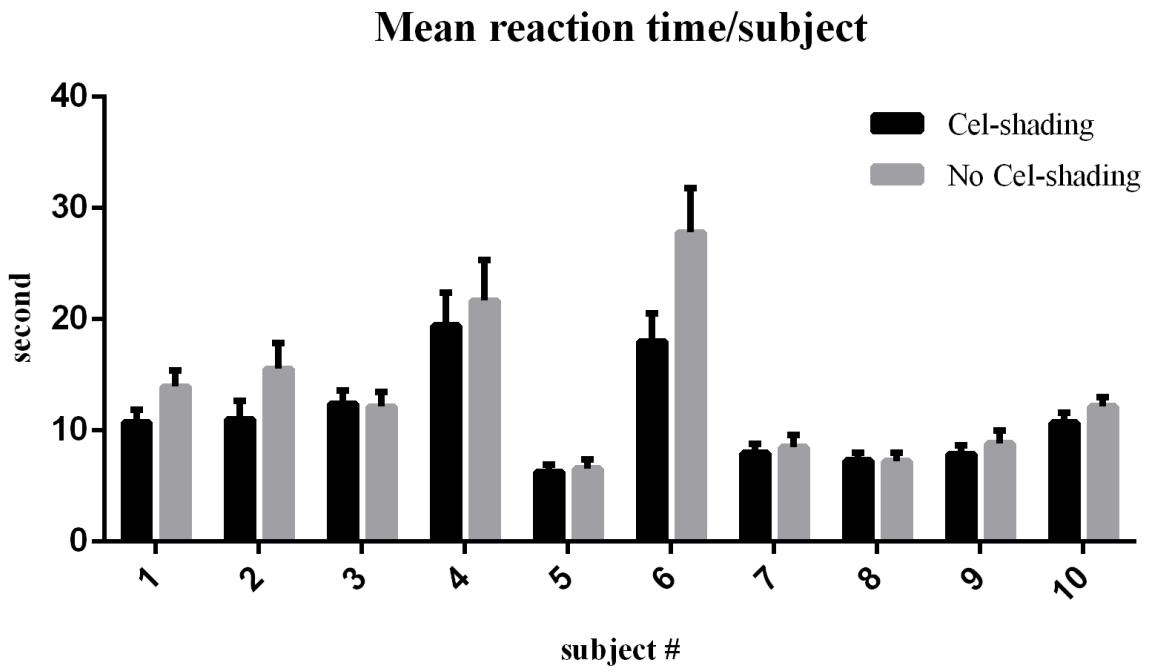


Figure 3.9: The subjects' response time (sec) with and without contour enhancement

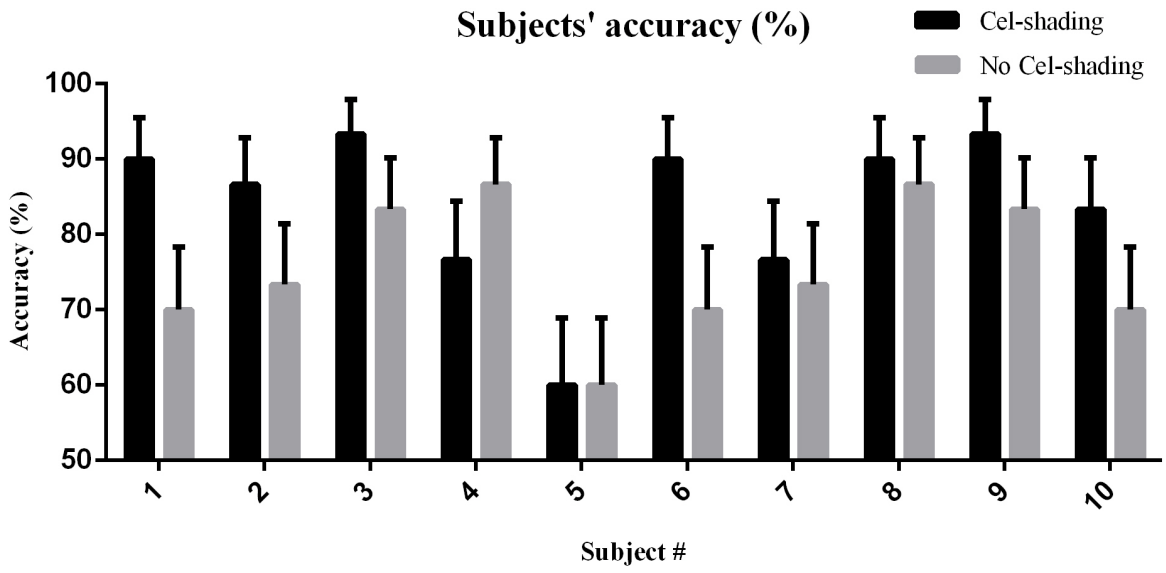


Figure 3.10: The subjects' accuracy (%) with and without contour enhancement

trials ($d' = (Z_{hit_{rate}} - Z_{FA_{rate}}) = 0.08 - (-0.96) = 1.04$) with the d' of non-enhanced trials ($d' = (Z_{hit_{rate}} - Z_{FA_{rate}}) = 0.04 - (-0.56) = 0.6$). A standard interpretation is that d' values greater than 1 reflect good recognition performance and values less than 1 represents low hit/FA ratio (d' at chance is always zero) [33, 34]. Thus, our results indicate better discrimination performance in enhanced trials compared to non-enhanced ones.

In addition to our objective measures, post-experiment interviews were conducted in which participants collectively agreed that tracing vessels was easier in the presence of contour enhancement.

3.4.2 Depth Perception Experiment

The subjects' overall performance and reaction time (RT) are shown in Table 3.2 and Figure 3.11. Our results illustrate the importance of including stereo and cel-shading contour enhancement for increasing the accuracy and/or lowering the RT for the range of difficulty of tasks. A significant increase in the overall accuracy (from 56% to 95.8%) coupled with a decrease in overall RT (from 4 sec to 2 sec) was observed when one moves from trials with *no contour enhancement & no stereo* to those with *full enhancement & high disparity*. Interestingly, in contrast to cel-shading darkness, stereo alone was found to have little to no significant effect on performance, particularly when the task is relatively difficult (Figure 3.11). This could be because visual cues of *occlusion* and *luminance* are strengthened when contours are enhanced, which in turn, provide a strong clue about the relative location of the cylindrical tubes even in the absence other cues; whereas stereopsis alone may not provide much information about the depth when the visual scene is featureless. In fact, the benefits of stereopsis have been shown to be task dependent [35] [36] [37]. In [14], for instance, stereopsis was shown to have no benefit in perceiving relative depth. In addition, subjects were informed that the level of cel-shading darkness corresponded to depth (section 3.2.1), and therefore could be used as an additional clue.

Although this pattern of improvement was repeated by further increasing the disparity or darkness, such increments did not always result in significantly improved outcome. For example, increasing the amount of darkness from *medium* to *full* or enabling stereopsis for objects with the lowest level of disparity did not produce statistically significant results ($p_{accuracy} = 0.4 > 0.05$ and $p_{RT} = 0.27 > 0.05$). In addition, full darkness around the contours might not be always desirable particularly in the presence of noisy data. Hence, improved performance can be achieved by providing cel-shading darkness just high enough to perceive the edge around the vascular structures but low enough to prevent exaggerating noise.

The results of our statistical tests under 12 possible combinations of stereopsis and contour enhancement (Table 3.1) are shown in Table 3.3 and Table 3.4). To control the type I error, Bonferroni correction was applied: let the significance threshold for each test be $p = 0.05$; then a stricter p-value based on Bonferroni correction would be: $\beta = 0.05/9 \approx 0.005$.

Disparity (Distance to Camera)	Level of darkness (Cel-shaded)					
	0 (None)		0.5 (Medium)		1 (Full)	
	RT (sec)	Acc. (%)	RT (sec)	Acc. (%)	RT (sec)	Acc. (%)
None	4.0±2.9	56.0±3.5	2.1±1.1	86.1±21.1	2.3±1.4	89.8±19.2
Low (566*cm*)	3.0±2.2	55.9±8.9	2.3±1.6	84.7±18.2	2.3±2.0	84.7±19.3
Middle (324*cm*)	3.1±2.1	67.3±10.2	2.7±2.4	91.6±11.7	2.1±1.5	90.9±17.9
High (118*cm*)	2.8±1.3	86.8±9.3	2.4±1.5	95.1±10.2	2.0±1.0	95.8±5.4

Table 3.2: Overall level of performance (RT and Accuracy) ($\mu \pm std$)

3.4.3 Qualitative Feedback

In addition to aforementioned experiments, a neurosurgery consultant with experience in assessing vascular malformations was asked to provide us with qualitative assessment after planning an AVM intervention with the use of our software and a 3D dis-

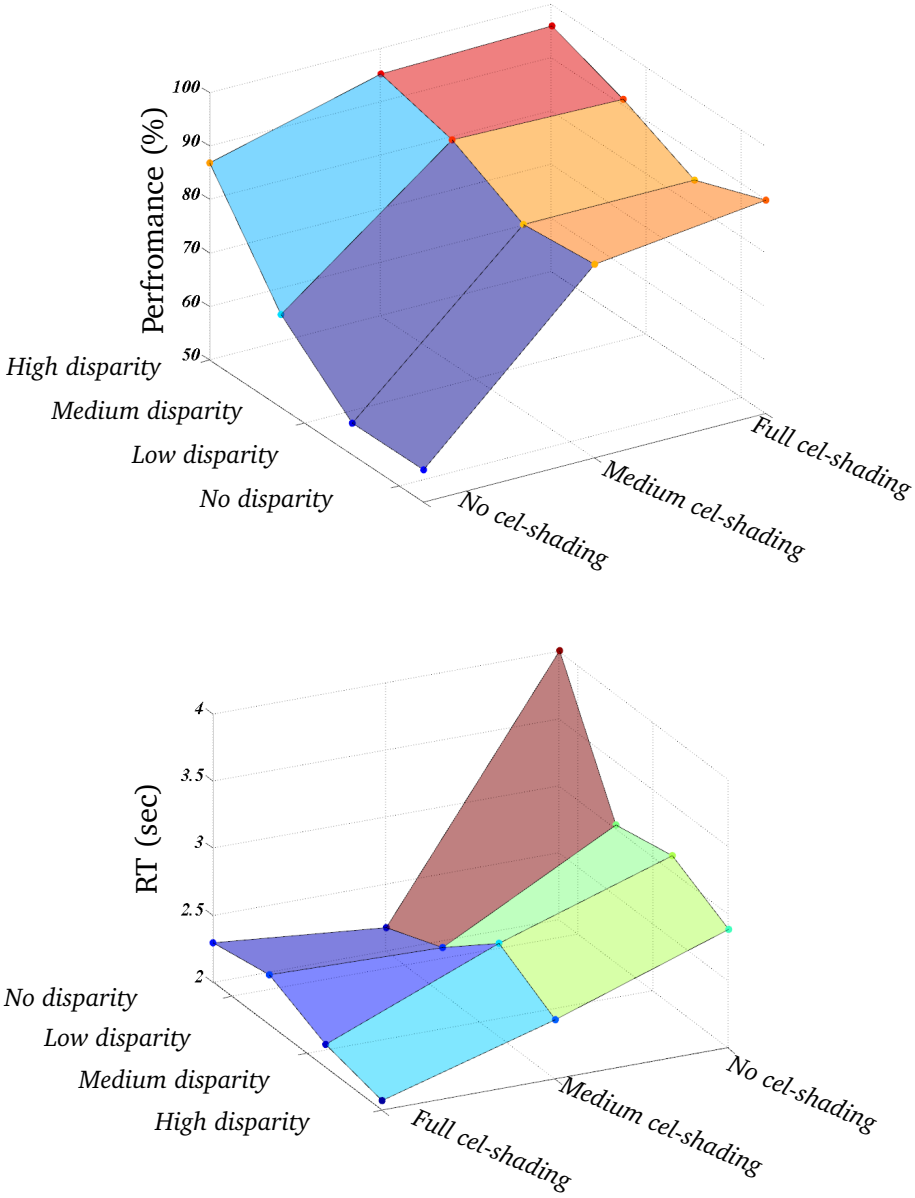


Figure 3.11: Mean accuracy (%) and mean RT (sec) (Refer to Table 3.2 for details)

	p-value	
Level of Darkness (Cel-shaded)	RT	Accuracy
between 0 (*None*) & 0.5 (*Medium*)	<0.005	<0.005
between 1 (*None*) & 0.5 (*Full*)	<0.005	<0.005
between 0.5 (*Medium*) & 1 (*Full*)	0.11	0.49

Table 3.3: Paired t-test comparison between subjects performance and RT under the possible combinations of contour enhancement

	p-value	
Disparity (Distance to Camera)	RT	Accuracy
between 0 ($NoStereo$) & 1.69 ($Stereo, Far$)	0.27	0.4
between 0 ($NoStereo$) & 1.80 ($Stereo, Middle$)	0.5	0.02
between 0 ($NoStereo$) & 1.85 ($Stereo, Near$)	0.17	<0.005
between 1.80 ($Stereo, Middle$) & 1.69 ($Stereo, Far$)	0.54	<0.005
between 1.85 ($Stereo, Near$) & 1.69 ($Stereo, Far$)	0.61	<0.005
between 1.85 ($Stereo, Near$) & 1.80 ($Stereo, Middle$)	0.34	<0.005

Table 3.4: Paired t-test comparison between subject performance and RT under the possible combinations of stereopsis

play[6]. Using the software, the expert could generate a volumetric representation of patient-specific data using 2D-TFs (Figure 3.12), adjust the shading models (Figures 3.13 and 3.4), and enable or disable the stereo view. Accordingly, as compared with the current standard-of-care, our interface was found more intuitive to:

1. locate the lesion of the nidus when stereo was enabled and the gradient-shading was applied;

2. perceive the continuity of the veins and arteries when stereo was enabled and both the gradient-shading and cel-shading were applied.

[6] 19" autostereoscopic TFT LCD with the resolution of 1280×1024; In this display, left/right images are reversible and can be displayed in full resolution.

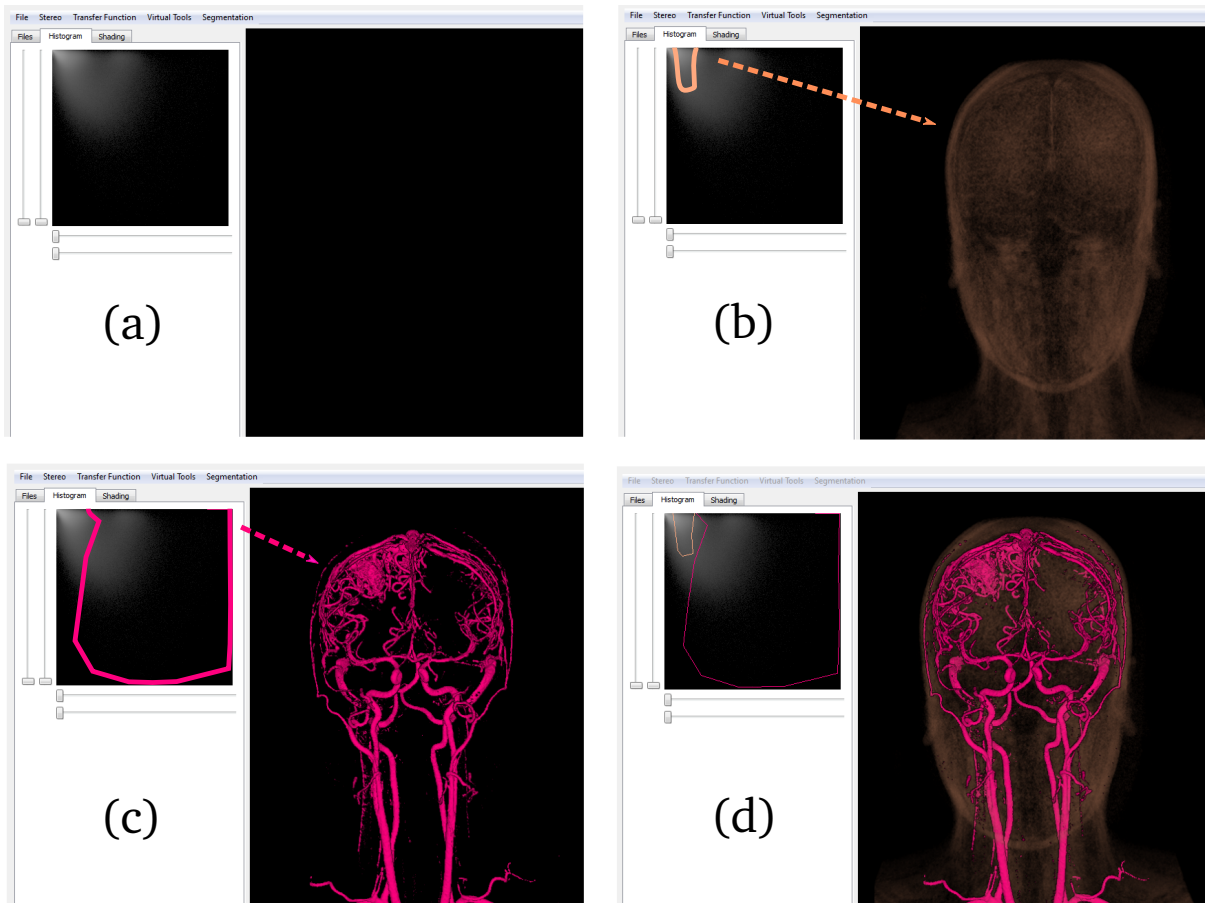


Figure 3.12: To generate a 2D-TF (*a*) a gradient-intensity histogram is generated automatically (top-left); (*b, c*) Different regions on the gradient-intensity histogram correspond to different tissue types (e.g., skin, vessels); (*d*) Users are equipped with tools to subjectively specify the region-of-interest on the histogram and set the colour/opacity accordingly.

Increasing the salience of cel-shading (> 50% of its maximum darkness), however, emphasized noise resulting in an undesirable outcome. Interestingly, such an increment in the amount of darkness was shown in section 3.4.2 to offer no significant performance benefit. It was also observed that defining the transfer function was a somewhat tedious task. Nevertheless, the expert surgeon recognized the novelty of contour enhancement and the advantage of using the aforementioned shading models coupled with stereopsis over conventional approaches for planning of AVM interventions.

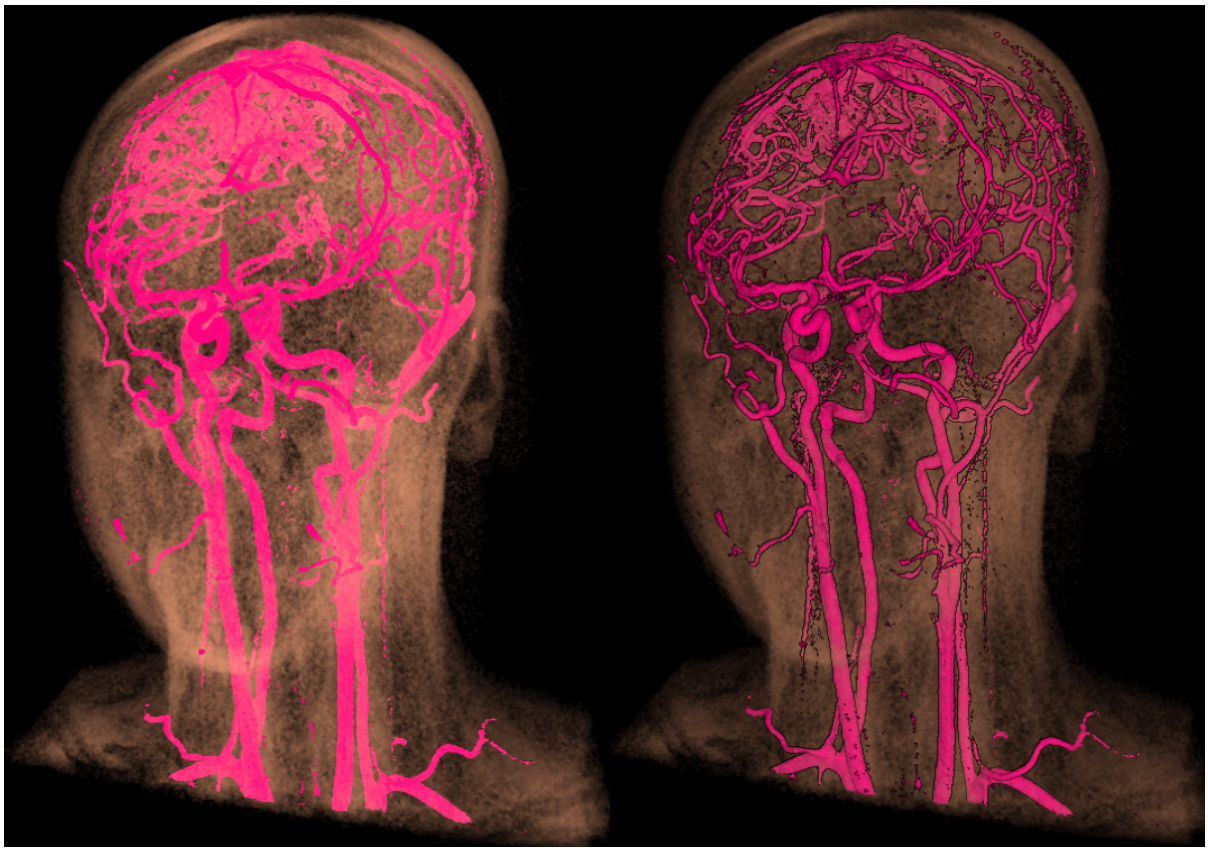


Figure 3.13: An AVM dataset before (left) and after (right) applying gradient shading and contour enhancement, augmenting the context of the volumetric visualization.

3.5 Conclusion

In current surgical practice, medical images play a fundamental role in the success or failure of an intervention with a significant impact on the patient outcome. To this effect, visualization is an important factor in preventing surgical errors and assisting surgeons with less experience in interpreting images with minimal effort. In fact, poor representation of medical images may lead to misperception and misdiagnosis, causing a large number of surgical accidents [38]. This notion is not limited to the operating room and can occur anywhere in the interventional workflow, from preoperative planning to postoperative assessment. In the case of AVM interventions, planning often involves the investigation of NC-MRA images in a series of axial-coronal-sagittal views to identify vessels-of-interest and their interspatial relationships. This task has been shown to be perceptually challenging, particularly for novice observers. To this end, we used techniques in computer graphics such as volume rendering and cel-shading to explicitly facilitate perception of vascular structures. Our results can be summarized with the following observations:

1. Enhancing contours results in improved understanding of *connectivity* relationships between different parts of a vascular structure;

2. Stereopsis and contour enhancement both facilitate *relative depth* perception;

3. Enhancing contours improves the perception of stereo. However, more enhancement may not always yield improved outcomes, and might even result in poorer performance in the case of noisy data.

In AVM interventions, such an improvement can potentially facilitate the process of planning and improve surgical outcomes by preventing perceptual errors and mitigating mental effort in the examination and exploration of preoperative images.

3.5.1 Limitations and Future Work

One major limitation of this study is that depth cues such as motion parallax, relative size, and aerial perspective were controlled out for purposes of scope. These cues can be enhanced or incorporated into visualization pipeline along with stereopsis and occlusion. Furthermore, although transfer functions are often predefined and immutable, the selection thereof plays a distinct role in visualization of medical images. Defining transfer functions, particularly for visualizing MR images, is a well-known challenge, but out of scope of this study. Other sources of limitations were small number of participants and lack of observation and characterization of learning effects. As future work, we would like to extend this work by using more clinically relevant stimuli, investigating the effect of depth cues other than contours and stereopsis on the perception of vascular structures, and expanding our subject population size to include both novices and expert neurosurgeons/neurologists.

Bibliography

[1] NINDS. Arteriovenous malformations and other vascular lesions of the central nervous system fact sheet, Feb. 2011.

[2] Stephen L Ondra, Henry Troupp, Eugene D George, and Karen Schwab. The natural history of symptomatic arteriovenous malformations of the brain: a 24-year follow-up assessment. *Journal of neurosurgery*, 73(3):387–391, 1990.

[3] NA Martin, R Khanna, C Doberstein, and J Bentson. Therapeutic embolization of arteriovenous malformations: the case for and against. *Clinical neurosurgery*, 46:295–318, 1999.

[4] R Tyler Frizzel and Wink S Fisher III. Cure, morbidity, and mortality associated with embolization of brain arteriovenous malformations: a review of 1246 patients in 32 series over a 35-year period. *Neurosurgery*, 37(6):1031–1040, 1995.

[5] Carlo Schaller. Arteriovenous malformations in functional areas of the brain. *Brain*, 123(1):190–192, 2000.

[6] Stefano Baldassi, Nicola Megna, and David C Burr. Visual clutter causes high-magnitude errors. *PLoS biology*, 4(3):e56, 2006.

[7] Elizabeth Bullitt, Stephen Aylward, Estrada J Bernard Jr, and Guido Gerig. Computer-assisted visualization of arteriovenous malformations on the home personal computer. *Neurosurgery*, 48(3):576–583, 2001.

[8] Florian Weiler, Christian Rieder, CA David, C Wald, and Horst K Hahn. Avm-explorer: Multi-volume visualization of vascular structures for planning of cerebral avm surgery. *Spring Eurograp*, 6:2–5, 2011.

[9] Sean Jy-Shyang Chen, Marta Kersten-Oertel, Simon Drouin, and D Louis Collins. Visualizing the path of blood flow in static vessel images for image guided surgery of cerebral arteriovenous malformations. In *Proc. of SPIE Vol*, volume 8316, pages 831630–1, 2012.

[10] Timo Ropinski, Frank Steinicke, and Klaus Hinrichs. Visually supporting depth perception in angiography imaging. In *Smart Graphics*, pages 93–104. Springer, 2006.

[11] Alark Joshi, Xiaoning Qian, Donald P Dione, Ketan Bulsara, Christopher Breuer, Albert J Sinusas, and Xenophon Papademetris. Effective visualization of complex vascular structures using a non-parametric vessel detection method. *Visualization and Computer Graphics, IEEE Transactions on*, 14(6):1603–1610, 2008.

[12] Felix Ritter, Christian Hansen, Volker Dicken, Olaf Konrad, Bernhard Preim, and H-O Peitgen. Real-time illustration of vascular structures. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):877–884, 2006.

[13] Christian Hansen, Jan Wieferich, Felix Ritter, Christian Rieder, and Heinz-Otto Peitgen. Illustrative visualization of 3d planning models for augmented reality in liver surgery. *International journal of computer assisted radiology and surgery*, 5(2):133–141, 2010.

[14] Marta Kersten-Oertel, S Chen, and D Collins. An evaluation of depth enhancing perceptual cues for vascular volume visualization in neurosurgery. 2013.

[15] Christopher G Healey. Perceptual techniques for scientific visualization. In *SIGGRAPH99 Course*. Citeseer, 1999.

[16] Anne Treisman. Preattentive processing in vision. *Computer vision, graphics, and image processing*, 31(2):156–177, 1985.

[17] ZW Pylyshyn and RA Eagleson. Developing a network model of multiple visual indexing. In *Investigative Ophthalmology & Visual Science*, volume 35, pages 2007–2007. Lippincott-Raven Publ 227 East Washington Sq, Philadelphia, Pa 19106, 1994.

[18] Colin Ware. *Information visualization: perception for design*. Elsevier, 2012.

[19] Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano. Measuring visual clutter. *Journal of Vision*, 7(2):17, 2007.

[20] Cathleen M Moore, Steven Yantis, and Barry Vaughan. Object-based visual selection: Evidence from perceptual completion. *Psychological Science*, 9(2):104–110, 1998.

[21] Joe Kniss, Gordon Kindlmann, and Charles Hansen. Interactive volume rendering using multi-dimensional transfer functions and direct manipulation widgets. In *Proceedings of the conference on Visualization'01*, pages 255–262. IEEE Computer Society, 2001.

[22] Veronika Šoltészová, Daniel Patel, Stefan Bruckner, and Ivan Viola. A multidirectional occlusion shading model for direct volume rendering. In *Computer Graphics Forum*, volume 29, pages 883–891. Wiley Online Library, 2010.

[23] Feng Dong, Gordon J Clapworthy, and Mel Krokos. Volume rendering of fine details within medical data. In *Proceedings of the conference on Visualization'01*, pages 387–394. IEEE Computer Society, 2001.

[24] James E Cutting. How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments, & Computers*, 29(1):27–36, 1997.

[25] David Marr and Tomaso Poggio. A computational theory of human stereo vision. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 204(1156):301–328, 1979.

[26] Ian P Howard and Brian J Rogers. *Binocular vision and stereopsis*. Oxford University Press, 1995.

[27] Jerold W Wallis, Tom R Miller, Charles A Lerner, and Eric C Kleerup. Three-dimensional display in nuclear medicine. *Medical Imaging, IEEE Transactions on*, 8(4):297–230, 1989.

[28] Elliot K Fishman, Derek R Ney, David G Heath, Frank M Corl, Karen M Horton, and Pamela T Johnson. Volume rendering versus maximum intensity projection in CT angiography: What works best, when, and why? *Radiographics*, 26(3):905–922, 2006.

[29] James T Kajiya. The rendering equation. In *ACM Siggraph Computer Graphics*, volume 20, pages 143–150. ACM, 1986.

[30] David S Immel, Michael F Cohen, and Donald P Greenberg. A radiosity method for non-diffuse environments. *ACM SIGGRAPH Computer Graphics*, 20(4):133–142, 1986.

[31] Walter H Ehrenstein and Addie Ehrenstein. Psychophysical methods. In *Modern techniques in neuroscience research*, pages 1211–1241. Springer, 1999.

[32] Jacob Cohen. *Statistical power analysis for the behavioral sciences*. Psychology Press, 1988.

[33] Wilson P Tanner Jr and John A Swets. A decision-making theory of visual detection. *Psychological review*, 61(6):401, 1954.

[34] Micah R Bregman, Aniruddh D Patel, and Timothy Q Gentner. Stimulus-dependent flexibility in non-human auditory pitch processing. *Cognition*, 122(1):51–60, 2012.

[35] Frank Tendick, Sunil Bhoyrul, and Lawrence W Way. Comparison of laparoscopic imaging systems and conditions using a knot-tying task. *Computer Aided Surgery*, 2(1):24–33, 1997.

[36] L Thomas, C Wickens, and IL Savoy. Effects of cdti display dimensionality and conflict geometry on conflict resolution performance. In *Proceedings of the 13th International Symposium on Aviation Psychology*, 2005.

[37] Yili Liu and Christopher D Wickens. Use of computer graphics and cluster analysis in aiding relational judgment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34(2):165–178, 1992.

[38] Lawrence W Way, Lygia Stewart, Walter Gantert, Kingsway Liu, Crystine M Lee, Karen Whang, and John G Hunter. Causes and prevention of laparoscopic bile duct injuries: analysis of 252 cases from a human factors and cognitive psychology perspective. *Annals of surgery*, 237(4):460, 2003.

Chapter 4

Training for Planning Tumour Resection: Augmented Reality and Human Factors

This chapter is adapted from '*Training for Planning Tumour Resection: Augmented Reality and Human Factors*'[1], '*Use of a Mixed-Reality System to Improve the Planning of Brain Tumour Resections: Preliminary Results*'[2], and '*The Role of Augmented Reality in Training the Planning of Brain Tumor Resection*'[3].

My contribution to this chapter involved (i) developing the AR/VR system, (ii) designing and conducting experiments, (iii) analyzing data, and (iv) writing manuscripts.

[1]Abhari, K., Baxter, J., Chen, E., Khan, A., Wedlake, C., Peters, T., de Ribaupierre, S., Eagleson, R. "Training for Planning Tumour Resection: Augmented Reality and Human Factors", Revised and Resubmitted, Submission No. TBME-00725-2014, *IEEE Transactions on Biomedical Engineering*, (2014)

[2]Abhari, K., Baxter, J., Chen, E., Khan, A., Wedlake, C., Peters, T., de Ribaupierre, S., Eagleson, R. "Use of a Mixed-Reality System to Improve the Planning of Brain Tumour Resections: Preliminary Results" Lecture Notes in Computer Science, Augmented Environments for Computer-Assisted Interventions, Springer Berlin Heidelberg, 7815: 55-66, (2013)

[3]Abhari, K., Baxter, J., Chen, E., Khan, A., Wedlake, C., Peters, T., de Ribaupierre, S., Eagleson, R. "The Role of Augmented Reality in Training the Planning of Brain Tumor Resection" Lecture Notes in Computer Science, Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions, Springer Berlin Heidelberg, 8090: 241-248, (2013)

4.1 Introduction

4.1.1 Clinical Motivation

In 2013 alone, 26,000 North Americans were diagnosed with brain cancer, resulting in 16,000 deaths [1, 2]. It is estimated that 680,000 Americans are affected by some form of primary brain or central nervous system tumours (20% of which are malignant) [3]. The five-year relative survival rate of brain cancer is only 25%~35%, making it one of the least survivable types of cancer in North America [1, 2]. Effective treatment is therefore necessary to prevent complications which can develop with brain cancer, ranging from loss of vision and speech to paralysis and death.

Among the different courses of treatment, surgical removal of the tumour is often recommended with the resection constrained to preserve the brain's healthy tissues and functional status. Surgery is also indicated to perform a biopsy or implant a radiation source or chemotherapeutic agent [4]. Furthermore, total removal of the tumour by surgery may be the only option in the case of many benign tumours [4]. The success of these interventions greatly depends on the accuracy of planning and navigating to the target tumour. The goal of preoperative planning is to reduce intra- and post-operative complications by minimizing damage to healthy tissues and eloquent brain structures, particularly in the case of deep-seated tumours where planning is more critical. This is often accomplished by determining the optimal point of entry, the extent of the craniotomy, and the surgical pathways through which to advance instruments for debulking and removing the tumour.

Traditionally, in the absence of any three-dimensional imaging to guide them, neurosurgeons learned to form a mental representation of the brain to understand the spatial relationship between the lesion and surrounding structures and landmarks. This representation is built up from their prior anatomical knowledge and experience while scrolling through a sequence of 2D orthogonal slices (Figure 4.1-a), or through inter-

action with 3D images of preoperative magnetic resonance (MR) or computed tomography (CT) scans presented on a computer display (Figure 4.1-c). The ease of use of such perceptual environments - as will be discussed in the following section - is heavily influenced by their mode of visualization and interaction. This is even more true in the case of junior residents, whose more limited experience affects their ability to quickly perform spatial reasoning. In the current standard of care, residents gradually acquire these skills over several years throughout their residency by observing expert neurosurgeons planning their approach. Meanwhile, they rely heavily on neuronavigation systems, which may help them decide the surgical approach, but are not designed to improve their spatial reasoning abilities.

4.1.2 Background

Interaction within an environment, from determining the position of the cursor on the screen to controlling an aircraft, involves movement in some coordinate space that can be estimated relative to a *frame of reference*. The choice of this frame is task-dependent and can be relative to the entity of interest (object frame), the self (egocentric frame), the environment (allocentric or environment frame), or a display (display frame) [5, 6]. To interact with the system, one must first establish a frame of reference, and then perform a series of *mental transformations*. These transformations correspond to mentally rotating, translating, or scaling an entity to transform its spatial location and orientation from one frame of reference to another. For example, consider the situation in which a surgeon (egocentric frame) is navigating a surgical tool (object frame) by observing the endoscopic video projected on a display (display frame) inside an operating room (environment frame). Accordingly, the surgeon is required to perceive, reason, and act in a way that accommodates the transformation of information from the display to the frame of reference of the patient. Performing these transformations requires mental resources that may cause *cognitive overload* and reduce performance if

the capacity of the working memory is exceeded [7]. In the realm of neurosurgery, the choice of display (and its underlying methods of *visualization* and *interaction*) is perhaps one of the most important, yet under-appreciated, factors in task performance. This is of particular importance in neurosurgical planning, in which the entire process can vary significantly depending on the modes of visualization and interaction. Conventionally, planning is performed by scrolling through 2D orthogonal slices (2D), or interacting with 2D crossed-planes (XP) or 3D volume or surface rendered images (3D) [8] (Figure 4.1-a:c).

In addition to these conventional approaches, our study explores the possibility of using a mixed *virtual-reality* (VR) and *augmented-reality* (AR) environment in surgical planning of tumour resection interventions, with a particular focus on AR. Unlike VR where the entire scene is computationally generated, AR is often described as an environment in which the view of the real world is enhanced by *overlaying* computer-generated information. One such approach is purposed to complement the available information with computer generated images, providing a rich view of the operative field and so facilitating the performance of a surgical task. An example of an AR environment is illustrated in Figure 4.1-d. AR is sometimes referred as *mixed-reality*, a concept described and popularized by Milgram et al. [9]. In their proposed reality-virtuality continuum (Figure 4.2), the mixed-reality spans from environments in which the virtual information augments the real view (AR) to those where real information augments the virtual scene (augmented-virtuality or AV).

4.1.3 Related Work

Since the introduction of immersive environments in 1957 [10], a body of work has been devoted to the development of virtual- and augmented-reality environments for surgical training, planning, and navigation (refer to [11, 12, 13, 14, 15] for more comprehensive reviews). In the realm of neurosurgery, a number of AR simulators have

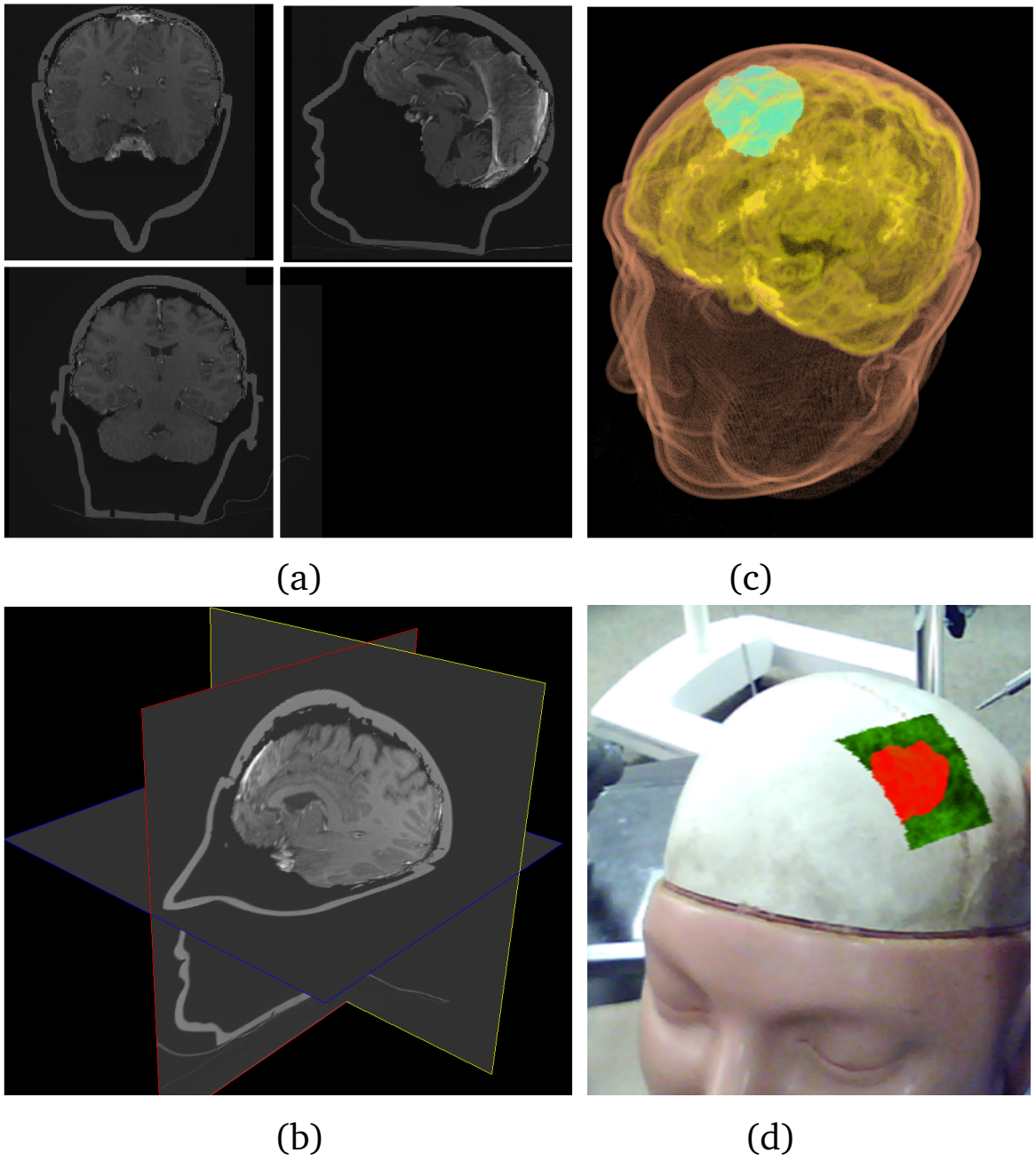


Figure 4.1: Planning environments: a) 2D views of axial/coronal/sagittal slices, b) crossed-plane (XP) representation of 2D slices, c) 3D volume rendering, d) Overlay of virtual images on the real video in an augmented-reality (AR) environment

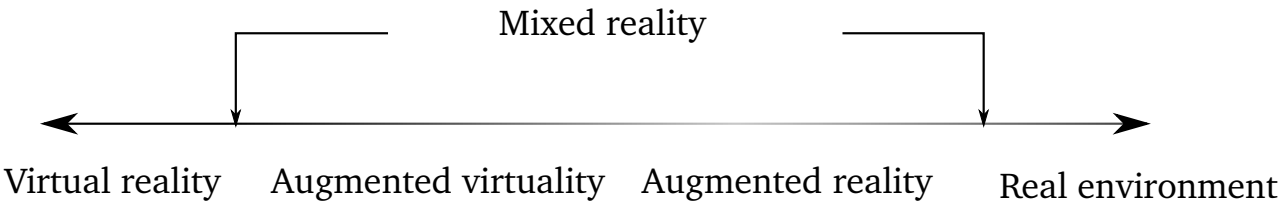


Figure 4.2: the reality-virtuality continuum proposed by Milgram et al. [9]

been developed to allow surgeons to plan and rehearse a surgical approach prior to the actual operation, with the *VR-workbench* [16], *Dextroscope*™[17], and *Immersive-Touch*® [18] platforms being a few examples. In these simulators, the user must wear a pair of active stereo glasses while holding a set of controllers positioned beneath a half-silvered mirror. By doing so, it is then possible to visualize and manipulate patient volumetric images simultaneously where the visual and haptic environments are registered. A similar concept was employed by Fichtinger et al. [19] to facilitate the needle insertion in a head phantom inside the CT scanner. Although these systems provide a safe training and planning environment, user interaction is limited due to restricted working space, small field of view, and sometimes absence of head tracking. A more natural method of interaction was proposed by Hinckley et al. [20], which involved holding a miniature size head mannequin in one hand and a plastic plate in the other hand to virtually slice-open the volumetric representation of patient data presented on a computer display. Although more intuitive, the interaction appeared unnatural whenever the operator's perspective and the virtual slice were misaligned.

AR also has been shown to facilitate intra-operative navigation in minimally-invasive neurosurgery. In [21], for instance, preoperative images of brain landmarks and critical structures were superimposed on a live video stream of the patient's head (or the surgical site in [22]) and displayed on a screen during surgery. The *Microscope-Assisted Guided Interventions* (MAGI) system [23] is another example where 3D preoperative images were overlaid into both eyepieces of the binocular optics of a surgical microscope. Shahidi et al. [24] made use of AR to improve the endoscopic view by reg-

istering real endoscopic and virtual preoperative images together. Bichlmeier et al. [25] proposed a so-called *virtual mirror* to display a desired perspective in endoscopic procedures when the virtual object cannot be viewed due to physical restrictions, a concept similar to the use of mouth mirrors in dentistry. In the *Reality Augmentation for Surgical Procedure* (RAMP) system [26], a tracked head-mounted display (HMD) was employed to display the neuroanatomical structures on a head phantom. In this system, a camera and an LED infrared flash is incorporated into the HMD, preventing the lag between rendering virtual and real images as well as the line of sight issue of the optical trackers. Lerotic et al. made use of a novel non-photorealistic technique for visualizing virtual object while maintaining salient anatomical details of the exposed surface [27]. Paul et al. [28] augmented intraoperative views of the operative field over preoperative images and projected the result onto a computer display. This compensated for the limited field of view, providing context to the intraoperative images. Similarly, Day et al. overlaid endoscopic images as they were acquired, onto 3D preoperative data in their correct anatomical positions [29].
With only modest exceptions (e.g. [30], [31]), the aforementioned systems suffer from lack of proper human factors evaluation and/or objective assessment.

4.1.4 Hypothesis, Rationale, and Objective

Hypothesis

We hypothesize that, compared to conventional planning environments, the proposed AR system offers a more intuitive visualization and interaction approach resulting in improved task performance to optimize surgical planning of tumour resection interventions.

Objective

In this study, we propose an AR system to assist novice surgeons in developing the cognitive skills demonstrated by experts in planning tumour resection interventions. This system can be used by trainees to assess and enhance their basic spatial skills and abilities, and eventually to plan an intervention. Furthermore, the capacities and constraints of the human visual system is considered in our visualization approach in order to improve the process of employing both virtual and real images. Experiments are conducted to evaluate the proposed system by measuring the user performance in both AR and conventional environments.

Rationale

Linte et al. [15] identified the lack of clinical evaluation and the difficulty of 3D information visualization and manipulation as two major barriers in introducing AR technology into clinical environments. Conducting user studies is therefore necessary to validate the benefit of AR environments compared to other approaches, and to provide insight towards perceptually-correct visualization methods [15]. This is mainly because non-intuitive modes of visualization and interaction, whether training, planning, or intra-operative navigation, are known to produce a significant cognitive load on surgeons [32], resulting in increased time or increased likelihood of errors under additional cognitive demands [6]. In contrast, intuitive training environments may reduce the learning curve, allowing novice surgeons to gain necessary visuospatial skills. There is a body of research on the trainability of such skills [33, 34, 35, 36].

4.1.5 Contributions

In this work, a novel virtual environment is designed, developed, and evaluated. The system is designed to provide an intuitive mode of interaction and visualization to

facilitate training neurosurgical visuospatial reasoning skills using multi-modal patient MRI images. This environment was evaluated based on clinical performance criteria relevant to brain tumour resection interventions.

4.2 Materials and Methods

In this section, a general framework for understanding the interaction and visualization aspects of planning environments in brain tumour resection is presented. This framework is used to guide the design, implementation, and evaluation of the AR environment described subsequently.

4.2.1 Action and Perception in Planning Environments

Different aspects of interactive systems, specifically interaction and visualization, can support different levels of visuospatial actions and perception. *Action* involves interacting with the system to update the visualization, which requires mental transformation from the self to the display frame of reference (e.g. moving the mouse to the right in order to rotate the image). *Perception*, on the other hand, involves observing the effects of the interaction and reconciling them, performing mental transformation to match the visualization with the desired 3D representation of the image, or validating that information estimated from one viewpoint is legitimate in another. The continual cycle of interaction with the system to perform a task is extremely flexible and adaptive, and aspects of action and perception are present in each of the underlying mental processes associated with interactive systems (Figure 4.3) [37].

4.2.2 Mental Processes in Planning Environments

Using an interactive system often involves executing a collection of mental processes. The hypothetical model shown in Figure 4.4 represents three primary underlying men-

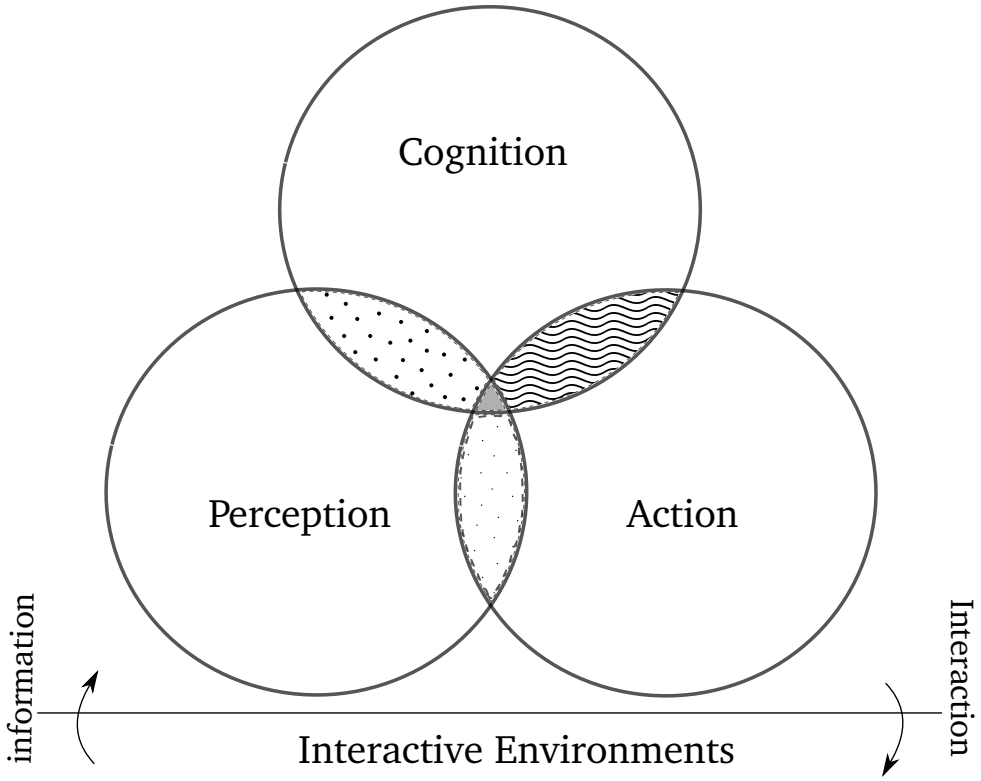


Figure 4.3: The perception-cognition-interaction cycle [37]

tal processes in determining the optimal point of entry and surgical path within a given planning environment. *Perceptual integration* involves assembling a 3D mental representation of the patient's brain by continuously obtaining and processing views. This representation is necessary for *action planning*, that is, mentally formulating and identifying the optimal surgical path and point of entry. The specific goals to be addressed are based on a number of criteria described in section 4.2.6 and may involve continuous mental processing of the information within a 3D space. Additionally, one must *coordinate between reference frames* to transform the mental representation of the optimal path/point of entry from the display frame of reference to the object frame of reference, i.e. the patient's skull in the operation room. These processes may occur simultaneously or sequentially depending on the preferences and experience of the user. Nevertheless, minimizing the mental demand elicited by these processes is necessary

to design an intuitive planning system.

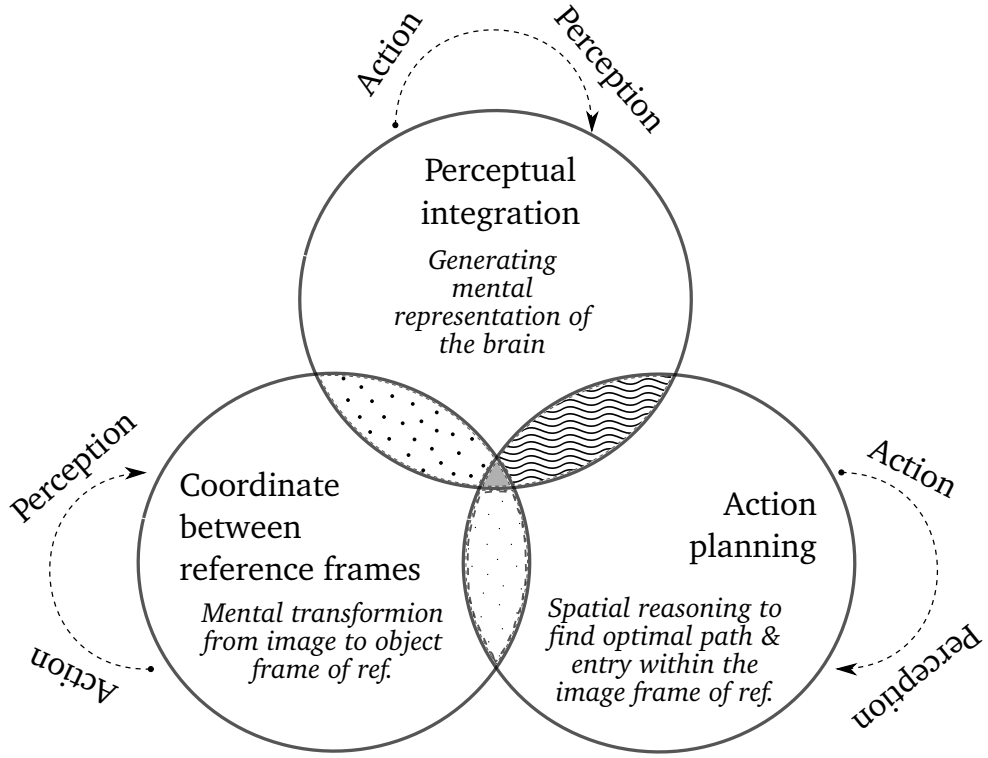


Figure 4.4: Underlying mental processes in planning environment

4.2.3 Mental Demands in Planning Environments

Perceptual integration in 2D involves generating a 3D mental representation of the brain by scrolling through 2D images, while recalling previous views. This process imposes a high demand on working memory and can be facilitated by positioning and interacting with these 2D images within a three-dimensional context (e.g. the XP environment). Visualizing the brain in 3D with the help of volume rendering can further lower the demand on spatial reasoning, reducing the need for perceptual integration (e.g. the AR and 3D environments) .

The *coordination of reference frames* required for transforming the formulated optimal path and point of entry into the physical frame of reference (e.g. patient's head or a head phantom) is almost negligible in the AR environment. This is because pre-

operative images are often overlaid directly on the physical object, eliminating the transformation between the two. In contrast, extensive mental work is needed in the case of 2D, XP, and 3D environments, particularly for novice operators, as they must transform the view from the display into the physical frame of reference.

Furthermore, exploring data via interaction with the phantom in AR environments mimics direct hand manipulation. Such a natural, everyday human behaviour can further reduce the mental demand of relevant judgment tasks in *action planning*. It also provides extra-retinal signals of different sources (such as proprioceptive information), enhancing the perception of depth when paired with certain visualization techniques (details are discussed in the next section). Although a more realistic representation of the brain in 3D facilitates the process of action planning compared to 2D and XP, less intuitive modes of interaction in these conventional environments demand a high degree of mental effort to formulate the surgical approach.

4.2.4 System Implementation

Our AR system comprises several modules as shown in the Figure 4.5. These modules can be broken down into five overarching categories:

1. The AR goggles and other tools,
2. Tracking and calibration components,
3. Medical image components,
4. AR visualization modules, and
5. VR visualization modules.

The AR goggles (Vuzix 920AR, Vuzix corporation, Rochester, NY) consist of a set of stereoscopic cameras (480×640) that provide a video feed into the planning system. In addition, the Vuzix goggles contain stereoscopic displays (600×800) on the opposite

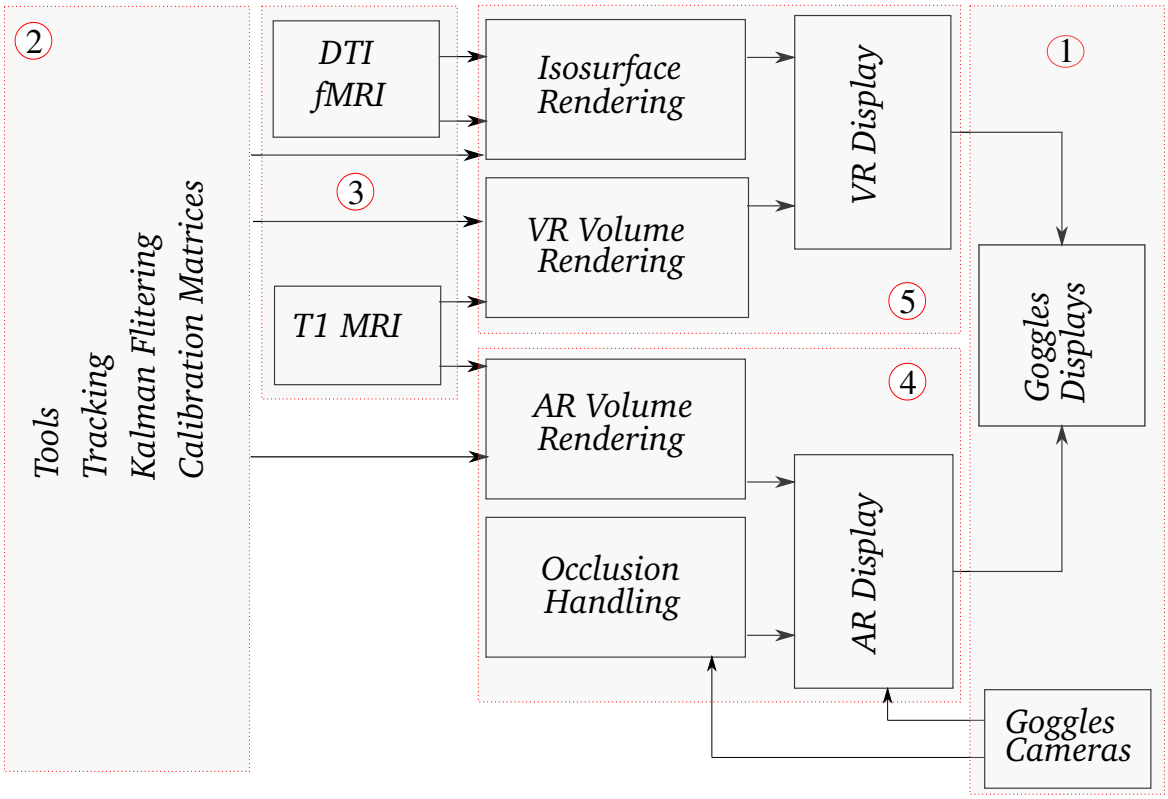


Figure 4.5: Implementation Diagram

side on which the system's renderings can be displayed. Other tools include a head phantom and a stylus, giving the user physical analogs for virtual representations of the patient and aperture respectively (Figure 4.6). Additionally, a set of foot pedals was incorporated to allow the user to interact with the system while maintaining their view-point of the phantom.

Tracking input to this system includes physical coordinates from an optical tracking system (Polaris, Northern Digital Inc., Waterloo, Canada), which relate to the AR glasses, tracked stylus, and head phantom through pre-acquired calibration matrices. The calibration process for the AR glasses is described in [38]. The tracked transform was smoothed using a Kalman filter framework to reduce jitter [39]. The state vector is the pose information, encoded using a 4×1 quaternion for rotation and 3×1 for translation. The process model is the identity matrix, and the measurement model is the actual measurement from the optical tracking system. This tracking information

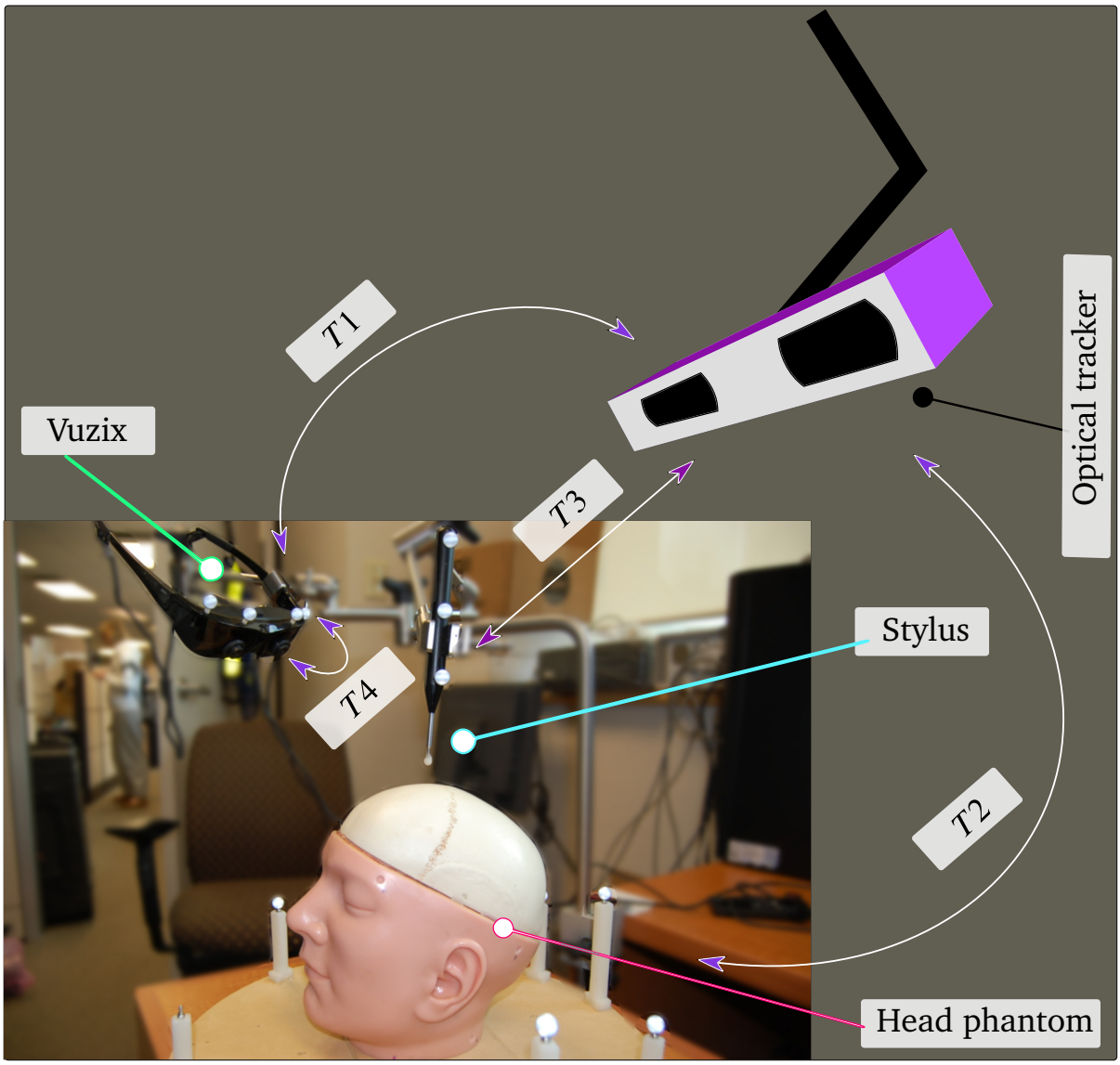


Figure 4.6: The AR system consists of a pair of AR goggles, a head phantom, a stylus, and an optical tracker. The transformation matrices of $T1$, $T2$, and $T3$ are given using the optical tracker while $T4$ is computed using camera calibration

is used to manage a virtual coordinate system in which the AR goggles and the head phantom dictate the position of the renderer viewpoint and the position of the medical image data respectively.

The system is capable of rendering T_1 weighted MRI images, using volume rendering integrated with a two-dimensional transfer function [40], and displaying associated fMRI activation clusters and white matter tracts. T_1 weighted acquisitions are clinically used for detecting brain tumours and planning brain tumour resection.

In VR mode, there are two separate renderers. First, the isosurface renderer converts a polyhedral representation of the DTI tracts, eloquent fMRI regions, and the stylus to a two-dimensional rendering of these objects. The depth buffer (z-buffer) for this rendering is then passed to the VR volume renderer, which uses ray casting and two-dimensional transfer functions to render the tumour and skin in the T_1 weighted MRI. This rendering style supports opacity values, and thus renders the background as transparent and the skin translucent, allowing internal structures to be visualized. These renderings are then merged together, overlaying the volume rendering on the isosurface rendering. The result is a three-dimensional rendering of the phantom with a layer of translucent skin through which the tumour, white matter tracts, and eloquent cortical regions can be seen. A virtual representation of the stylus is incorporated that corresponds to its location relative to the head phantom.

In AR mode, the volume rendering acts similarly, but is sensitive to a virtual aperture at the stylus tip. The same ray casting technique is used, but rays can terminate prematurely if they do not pass through the aperture. This conveys the appearance of a *window* into the head phantom at the tip of the stylus, which is sensitive to the position and orientation of the Vuzix glasses relative to the phantom. The occlusion handling system is designed to identify regions of the video feeds that are occluded by the participant's hands (Figure 4.7[4]). This is accomplished via hue-based thresholding which

[4]In this paper, the stereoscopic views are illustrated in left-right-left fashion, corresponding to left-right-left AR cameras and displays.

masks out the blue surgical gloves worn by the participant (refer to 4.2.5 for more details). The final step in the AR display is to merge the AR volume rendering with the occlusion sensitive video achieved by masking out occluded areas in the AR volume rendering, marking them as transparent. Subsequently, the AR volume rendering is overlaid on the AR video feed to the goggles. Furthermore, as part of our visualization approach, users have an option to substitute the virtual window with a set of grid lines (Figure 4.8) to improve the perception of depth (refer to 4.2.5 for more details).

The desired rendering system is selected by the user; a foot pedal can be used to toggle between the AR and VR display modes. This foot pedal also can be used to lock the position of the stylus, allowing for the aperture to remain locked in position, giving the user the opportunity to use both hands to manipulate the head phantom. The tracking status is provided at the bottom of the visualization, indicating when a tracked object has valid line-of-sight with the tracker (green), has lost tracking information (red), or has its position locked (blue).

4.2.5 Perceptual Cues and Considerations in AR

While effective visualizations can dramatically improve the outcome and efficiency of interactive environments, poor visualization approaches may result in cognitive overload or misinterpretation [41]. This is even more profound in AR environments, as subtle conflicts between real and virtual images result in incorrect depth perception, degrading performance [13]. Although many studies have discussed the issue of depth perception in AR systems [42, 43], problems still persist [13]. In HMD-based AR systems, for instance, *accommodation-vergence conflict* may cause depth distortion, hindering performance [44]. Accommodation is an involuntary movement of the eye to keep objects in focus by changing the shape of the lens. Vergence is the inward/outward movement of both eyes in order to fixate on a single point. Normally, these two eye movements act in synchrony, providing coherent cues for the perception of depth. In

HMD-based systems, however, the observer's eyes accommodate to the screens while constantly diverging and converging to maintain the binocular vision on a single object, providing incoherent depth cues. This has been shown to be a major drawback of HMD-based AR systems [42]. Such conflicts, however, can be mitigated if more dominant visual cues convey the desired sense of depth. This section is therefore devoted to describe the visualization techniques used in the AR system to improve the perception through the most effective depth cues [45] including *stereopsis*, *motion parallax*, and *occlusion*.

Cel-shading

In chapter 3, contour enhancement is shown to facilitate the perception of relative depth [46][47]. This is even more profound in HMD-based ARs as the effect of contour enhancement is accentuated when paired with stereopsis [47].

Thus, in this work, cel-shading [47] (refer to chapter 3 for more information) is employed to enhance the contours of the target tumour (Figure 3.13).

Occlusion Handling

Partial blockage of one object's view by another object or *occlusion*, is the strongest cue in perceiving the relative proximity of objects [45]. One well-known problem with the use of AR environments for medical applications is the occlusion of operators' hands or medical devices by virtual images, resulting in depth misperception. To resolve this issue, hue-base thresholding was employed to detect and mask the blue surgical gloves worn by the participant (Figure 4.7). This *occlusion handling* technique was inspired by the work described in [48].

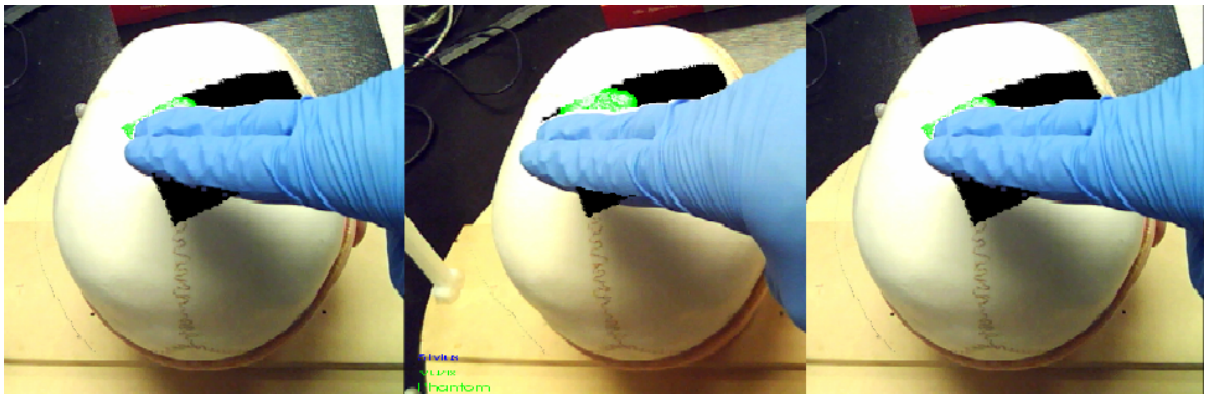


Figure 4.7: An example of hue-based occlusion handling

Grid lines

The apparent displacement of objects against their background, or *motion parallax*, is an effective and unambiguous cue, providing rich information about relative depth. In our AR environment, motion parallax occurs both when user makes lateral head movement with respect to the phantom (subject-motion), and when the phantom translates relative to the user (object-motion). Studies have shown that the former type of motion provides stronger sense of depth [49, 50], perhaps due to the fact that interaction with the phantom provides non-visual information such as proprioceptive, enhancing the perception in concert with motion parallax [51, 52]. Regardless of source of motion, a stronger sense of parallax can be evoked if the apparent displacement of the tumour is accentuated relative to its background. To this aim, a set of grid lines was placed behind the target tumour, amplifying the motional difference between the tumour and its background (Figures 4.8, 4.9).

Keyhole

Williams et al. [53] have shown that performing tasks that demand a high degree of visual attention lowers the accuracy of peripheral vision from 75% to 36%, shrinking the field of view. Therefore, when a visuospatial task requires focused attention (i.e. high cognitive consumption), it is undesirable to present a large amount of data in the

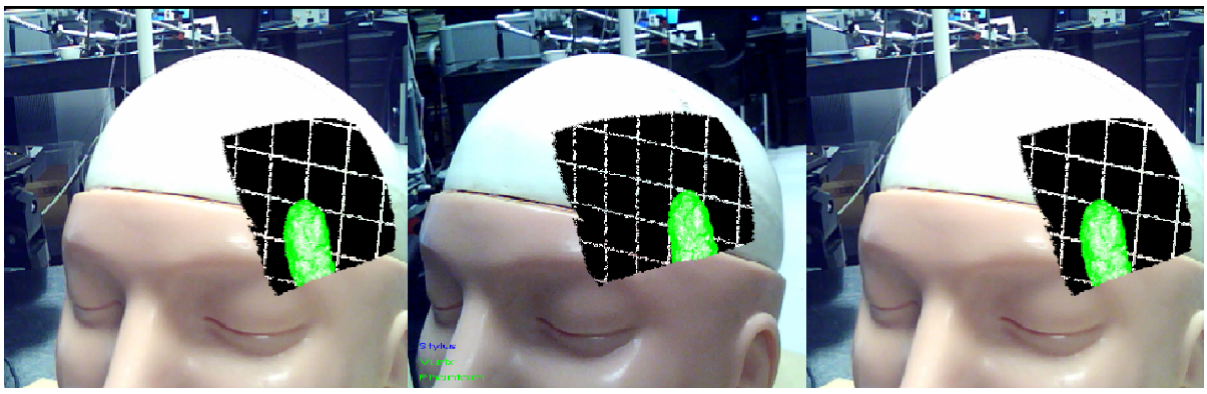


Figure 4.8: An example of use of a keyhole with grid lines to promote the sensation of depth

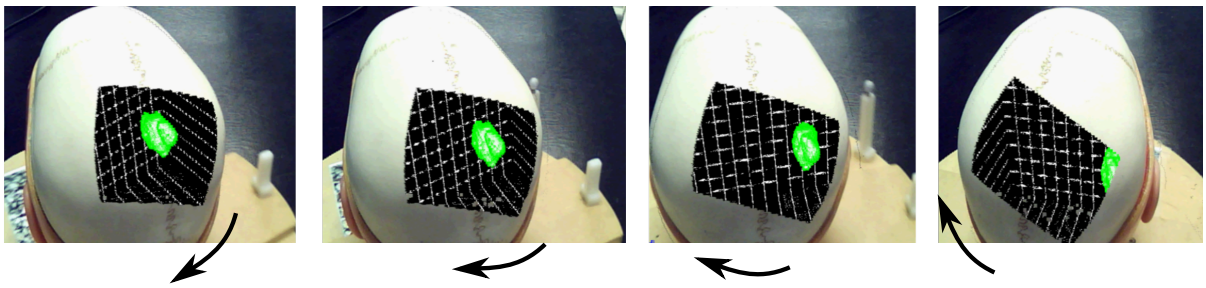


Figure 4.9: Visualizing grid lines behind the tumour evokes a strong sense of motion parallax while interacting with the phantom

periphery. This issue is addressed in [54] and [55] by presenting information through a virtual window. Similarly, in the proposed AR system, virtual images were placed inside the phantom to present the inner view within a *keyhole* (Figure 4.8).

Virtual reality

In AR environments, new information is superimposed on the real scene. This can extend the information to the user, but can also add clutter to an already complicated view, causing *information overload* [56]. Additionally, clutter in 3D can lead to visual occlusion of the target, degrading perceptual performance [57]. Accordingly, the amount of information presented in our proposed AR system was restricted to the target tumour and the keyhole. To visualize additional information, however, a virtual reality (VR) display was employed, in which the real world-view is removed by turning

off the video feed and halting the early ray termination process. In the VR mode, the user can visualize DTI tracts, functional areas of the brain, and the stylus' trajectory with no clutter (Figure 4.10). Users may benefit from both the AR and VR by toggling between these two modes of visualization.

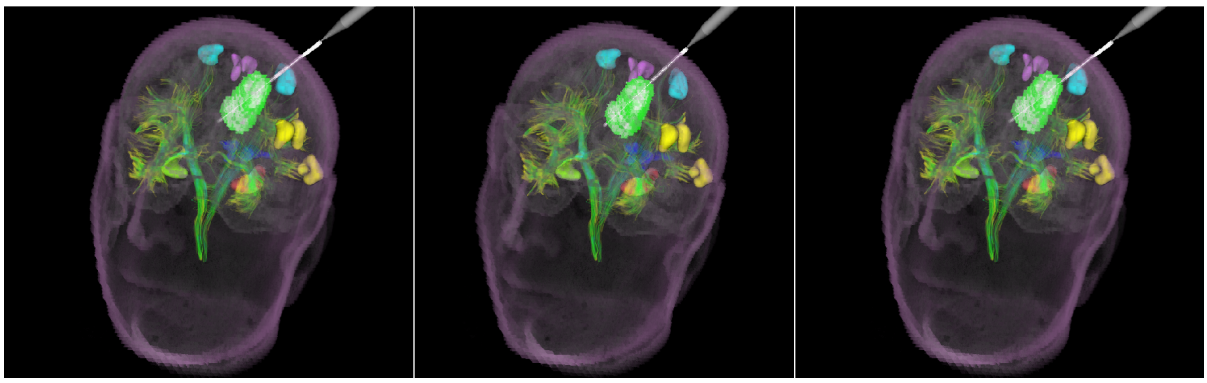


Figure 4.10: An example of the virtual reality mode to visualize DTI tracts, functional areas, and virtual stylus

4.2.6 Evaluation Studies

As discussed previously, the process of planning usually requires spatial reasoning in order to interact with preoperative images and transform the mental representation of the chosen path and entry point into the patient and operation environment. While 2D displays are currently considered the de facto standard for visualization and interaction with preoperative data, AR environments are believed to provide a more intuitive alternative. Although many research groups have proposed different AR environments for medical interventions (section 4.1.3), evaluation studies are necessary to understand the system's shortcomings and facilitate the translation into clinical practice. This motivated us to assess the effectiveness and efficiency of the AR system against conventional planning environments[5], exploring their underlying mental load (preliminary results have been presented in [58] and [59]).

[5]Neither occlusion handling nor grid lines were incorporated into the AR system for these evaluation studies

Our experiments involve three different phases. In phases 1 and 2, novice users interacted with synthetic images, whereas phase 3 involved clinicians and novice users performing clinically-relevant tasks by examining patient-specific data. These experiments were designed to address three major planning criteria:

- determining the shortest possible distance to the target tumour,
- aligning the surgical tool with the longest axis of the tumour, and
- specifying a trajectory that would avoid critical functional areas and white matter DTI tracts.

These criteria are the main considerations in planning tumour resection to minimize damage to healthy brain tissues. Our hypothesis was that AR can facilitate these clinically-relevant tasks significantly, improving the performance in specifying the surgical trajectory and point of entry.

Phase 1 **&** 2

Data

Stimuli consisted of the CT images of the head phantom with synthetic structures analogous to patient anatomical data (e.g. an ellipsoid representing the target tumour). The use of CT images with synthetically created anatomical structures allowed for more control over the experimental design.

Methodology

Each experiment involved 12 trials per subject (3[tasks] $\times$4[environments]) through which participants (8 male/2 female, no prior training) were presented with a randomized collection of synthetic phantom images. Each trial isolated one criterion by

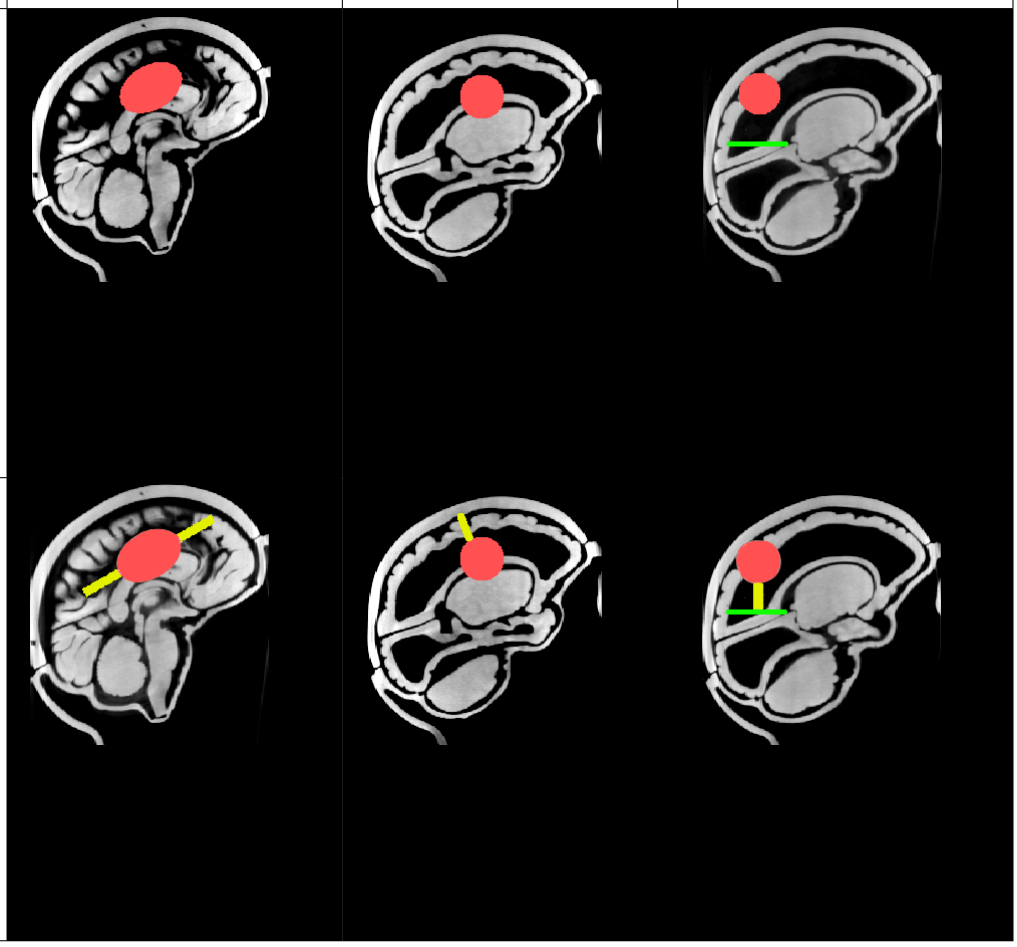
Task	LA	SD	AV
Description	*finding the longest axis of target*	*finding the shortest distance from target to skull*	*finding the max. Hausdorff distance between target & adjacent tube*
e.g. in 2D, phase 1			
e.g. in 2D, phase 2			

Table 4.1: *Middle row*: CT images of the phantom with a synthetic ellipsoid (LA), a sphere (SD), or a sphere and a tube (AV) were generated to be used as stimuli in phase 1; *bottom row*: The true longest axis of the ellipsoid (LA), shortest distance from the sphere to the skull (SD), and maximal Hausdorff distance from the sphere to the tube (AV) were shown in phase 2 (images are only a few examples from the 2D environment)

defining the task as either determining the *longest axis* of the tumour or the *shortest distance* from the skull to the tumour or the maximal distance from the tumour to a critical structure and so avoiding it. The ground truth for the longest axis (LA), shortest distance (SD), and avoidance (AV) tasks were determined, respectively, by computing the primary eigenvector of the tumour using principle component analysis, segmenting

the skull and computing its distance to the tumour, and finding the maximum Hausdorff distance between the tumour and the critical structure (Table 4.1). Subjects were asked, first, to investigate the data to estimate the optimal point of entry and surgical path, and then place and orient the stylus on the head phantom, disclosing their chosen location and orientation.

The rationale behind the second phase was to examine whether AR environments can facilitate the planning process, even if one makes use of a computer to assist with the estimation of the optimal point of entry and surgical path. Therefore, unlike phase 1, the ground truth, whether the shortest distance, longest axis, or maximal distance, was illustrated via synthetic lines (Table 4.1). Such visual assistance eliminated the need to formulate the optimal path and entry point, reducing the cognitive load mainly to the mental transformation from a display frame of reference (the computer's screen or HMD's display) to the object frame of reference (head phantom).

In both phases, the stimuli, the planning environment, and the task were randomized and counterbalanced to minimize the effects of learning and fatigue.

Analysis

As illustrated in Figure 4.11, user performance was measured based on the *translational error*, that is the Euclidean distance between the optimal points of entry and those selected by users (in mm), and the *rotational error*, that is the deviation between the optimal surgical path and the angle selected by participants (in degrees).

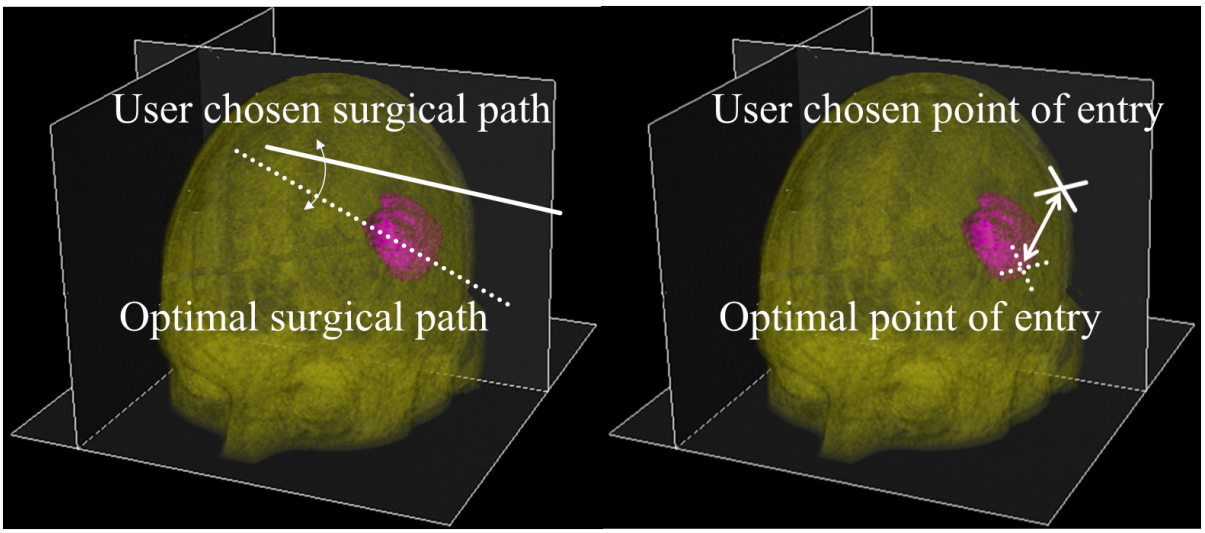


Figure 4.11: Rotational (*left*) and translational error (*right*) were used as metrics to measure users' performance

Phase 3

Data

Our medical image input comes from the MICCAI DTI challenge workshop[6] to generate a total of 112 different patient-specific cases by resizing and relocating the segmented tumour samples. All images were processed *a priori* to segment the tumour, relevant eloquent areas, and white matter tracts. This process was performed by an expert to create clinically relevant scenarios, and also to control the proximity of the tumour to surrounding eloquent areas. Such regions include functional areas of visual cortex, hippocampi, and areas representing language and the peripheral limbs, as well as the tractography of corticospinal, uncinate fasciculus, arcuate fasciculus, and the Meyer's loop[7] (Figure 4.12). Including DTI tracts and pseudo-fMRI data gives context to the experiments, increasing the ecological validity. For the purposes of this study, the MRI

[6]Medical Image Computing and Computer Intervention (MICCAI) 2010-11, permission is granted

[7]The *uncinate fasciculus* connects the orbitofrontal cortex to the temporal lobe and limbic system and is associated with emotion regulation, cognition, declarative memory and face recognition/memory. The Meyer's loop is essentially the prolongation of the optic tracts essential for vision. The arcuate fasciculus connects the Wernicke and Broca language areas. The corticospinal tracts is the descending fibers from the motor areas to the spinal cord enabling movements in the face, trunk and limbs.

images were registered and visualized within a CT scan of a head phantom.

Methodology

Each experiment involved 64 trials per subject[8] (2[tasks] ×4[environments] ×8[repeated trials]) in which 21 participants were asked to either identify the longest axis of the tumour (LA) and align the stylus with their chosen axis or determine the shortest distance to the tumour (SD) and place the stylus on their chosen location over the head phantom. The pool of subjects included 7 experts (3 neurosurgeons with >5 years of practice, 3 residents and 1 fellow with >2 years of training) and 14 novice graduate students.

Analysis

Rotational error for LA (longest axis) and translational error for SD (shortest distance) were measured after excluding inaccessible points of entry such as face, ears, and neck. In addition to accuracy metrics, the response time (RT) was also recorded, indicating the overall time one took to explore the images, formulate the optimal points of entry or surgical paths, and transform the results to the phantom frame of reference.

Index of performance: Because our task has a complicated three-dimensional structure, we redefined different aspects of Fitts' Law [60, 61]. First, we express the index of difficulty as:

$$ID = H[P(x)] - H[P(x|x \in A)] \tag{4.1}$$

where $H[\cdot]$ is the entropy functional, $P(x)$ is the uniform distribution over possible stylus tips/trajectories, and $P(x|x \in A)$ is the uniform distribution over possible stylus tips/trajectories that meet the criterion for success. Moreover, our methodology was an extension of Fitts' classical *click in the box* task [60]: Subjects were asked to specify their position and angles as quickly and accurately as possible, but there was no

[8] 48 trials for one of the experts

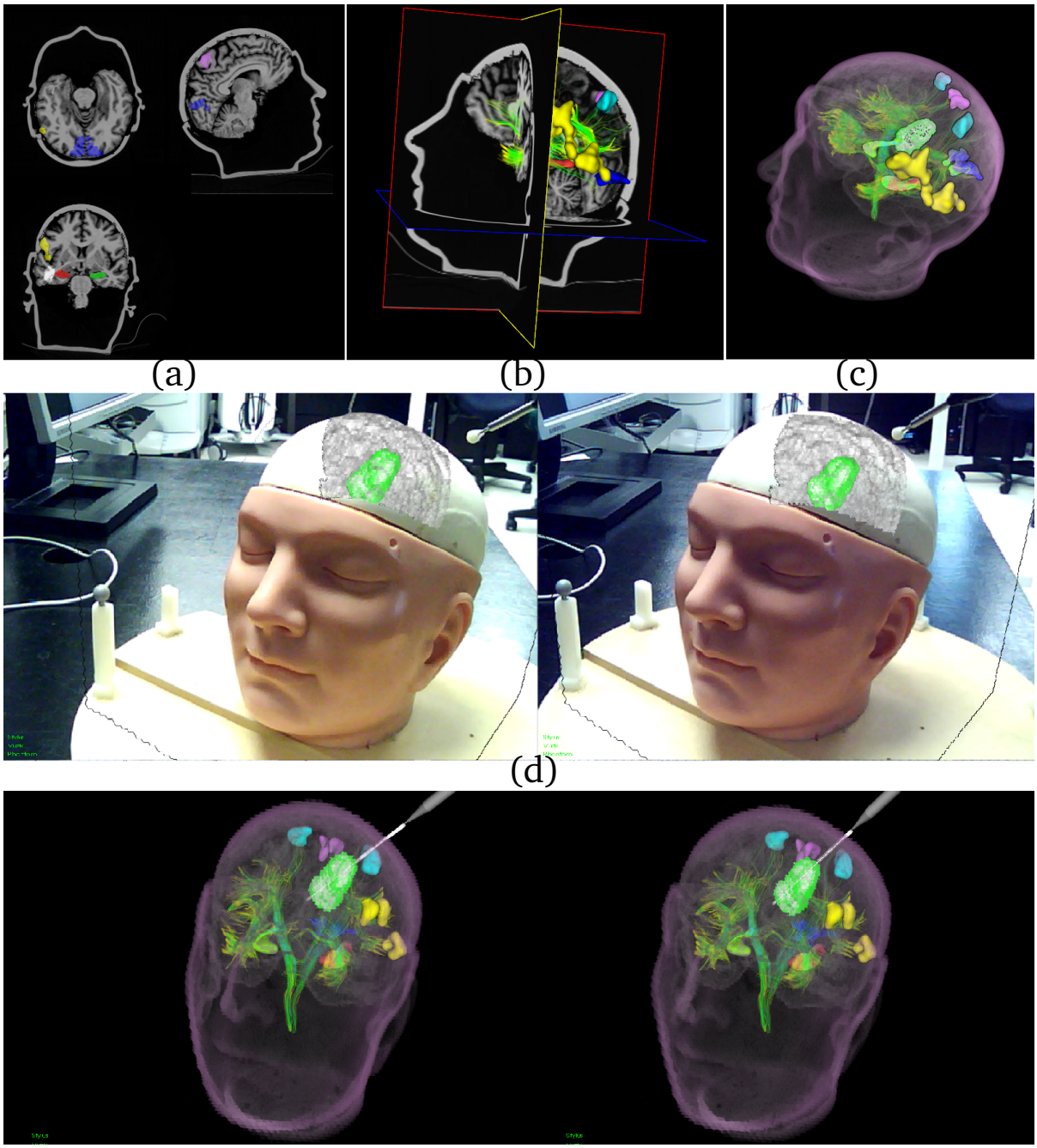


Figure 4.12: Visualization of patient-specific data in a) 2D, b) XP, c) 3D, and d) AR/VR (left and right images correspond to left and right views of the AR cameras/displays) environments in phase 3

pre-specified region that would correspond to a hit or miss within a region. Instead, the *effective width* of their targeting was derived from the mean of their position and angular responses. We thereby define a pseudo-criterion for each participant, specifically that $x \in A$ if and only if the error associated with x is less than or equal to the participant's average error. Accordingly, we can define the index of performance for each user as:

$$I_P = \frac{1}{T} ID = \frac{1}{T} \left(H[P(x)] - H[P(x|x \in A)] \right) \tag{4.2}$$

For the rotational error used in aligning the surgical trajectory with the longest axis, the index of performance can be determined analytically from the average angular error, μ (refer to the appendix for more details):

$$I_P = \frac{1}{T} \log_2 \left(\frac{1}{1 - \cos(\mu)} \right) . \tag{4.3}$$

For translation, the value of $H[P(x)] - H[P(x|x \in A)]$ is determined via a Monte-Carlo simulation similar to that employed to determine the gold standard entry point (refer to the appendix for more details).

4.3 Results and Discussions

4.3.1 Phase 1

The overall performance was calculated by computing the rotational and translational error (Figure 4.13). Since our hypothesis, i.e. the AR system lowers *both* rotational and translational error, may be interpreted methodologically as two independent hypotheses, the Šidàk correction was applied to control the type I error. Let the significance threshold for each test be $p = 0.01$; then a stricter p-value based on Šidàk correction would be: $\beta = 1 - \sqrt{1 - 0.01} \approx 0.005$, leading to a combined level of significance of 1%.

A two-way multivariate ANOVA indicated that the environment had a significant

effect on accuracy (rotational/translational error: $p<0.005$). Post-hoc analysis using Tukey HSD test revealed that users performed significantly more accurately in AR and 3D environments, indicating their lower rate of mental processes compared to XP or 2D environments. Furthermore, no interaction was observed ($p>0.05$) between the task and the environment. Furthermore, no significant difference was observed between the visualization environments in terms of the task completion time (overall time (sec): 178.8 ± 131.1 [2D], 204.1 ± 172.7 [XP], 143.6 ± 106.8 [3D], 170.1 ± 104.6 [AR]).

4.3.2 Phase 2

In phase 2, the task was restricted to a response that was dependent on *coordinating between different reference frames* to transform images from the display to the physical context. In this phase, AR was significantly superior to the other planning environments, reducing user dependent error. Both translational and rotational errors were significantly lower in AR compared to other environments ($p<0.005$, Figure 4.13). Unlike phase 1, interaction analysis revealed that the level of performance within the 2D and XP environments depended in part upon the task performed. Moreover, no significant difference was observed between the visualization environments in terms of the task completion time (overall time (sec): 111.5 ± 83.8 [2D], 123.6 ± 74.8 [XP], 78.7 ± 46.6 [3D], 77.2 ± 52.7 [AR]).

4.3.3 Phase 3

Phase 3 involved clinicians and novice users performing clinically-relevant tasks by examining patient-specific data. The result of this phase is depicted in Figure 4.14. A two-way multivariate ANOVA indicated that RT was significantly affected by both *expertise* and *environment* ($p<0.005$). The choice of *environment* also significantly affected the rotational error as well as I_p for both tasks ($p<0.05$). Regardless of expertise, the mean

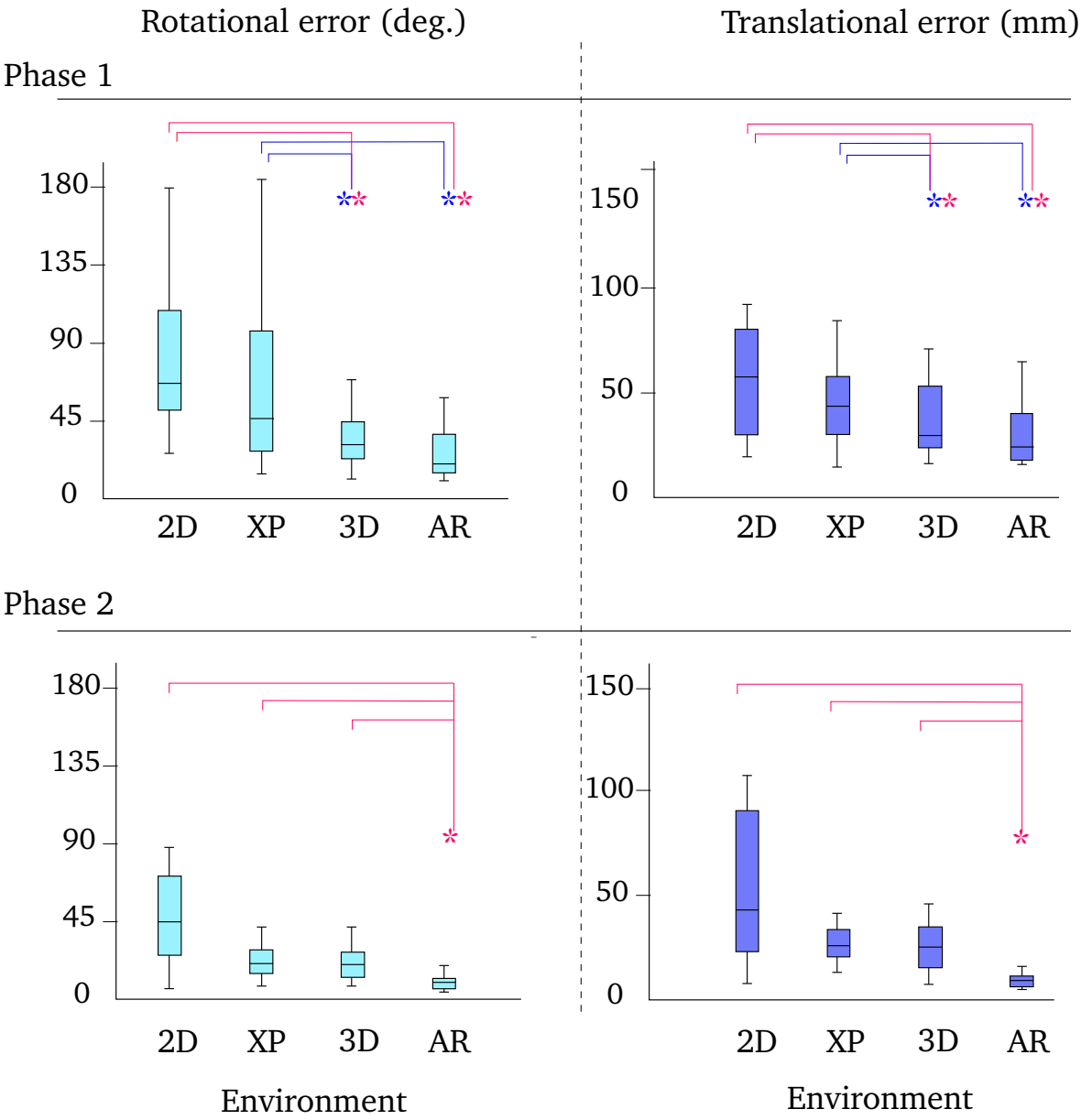


Figure 4.13: Overall rotational and translational errors observed in phase 1 (top) and phase 2 (bottom).

accuracy was higher in AR compared to other environments (Table 4.2), however, no significant difference was observed between these environments, perhaps due to insufficient power as a result of small sample size, or because of longer experimentation time which when coupled with poor ergonomics of the AR goggles resulted in fatigue and higher rate of error.

Unlike SD experiments, significant interaction was observed in LA ($p<0.05$), indicat-

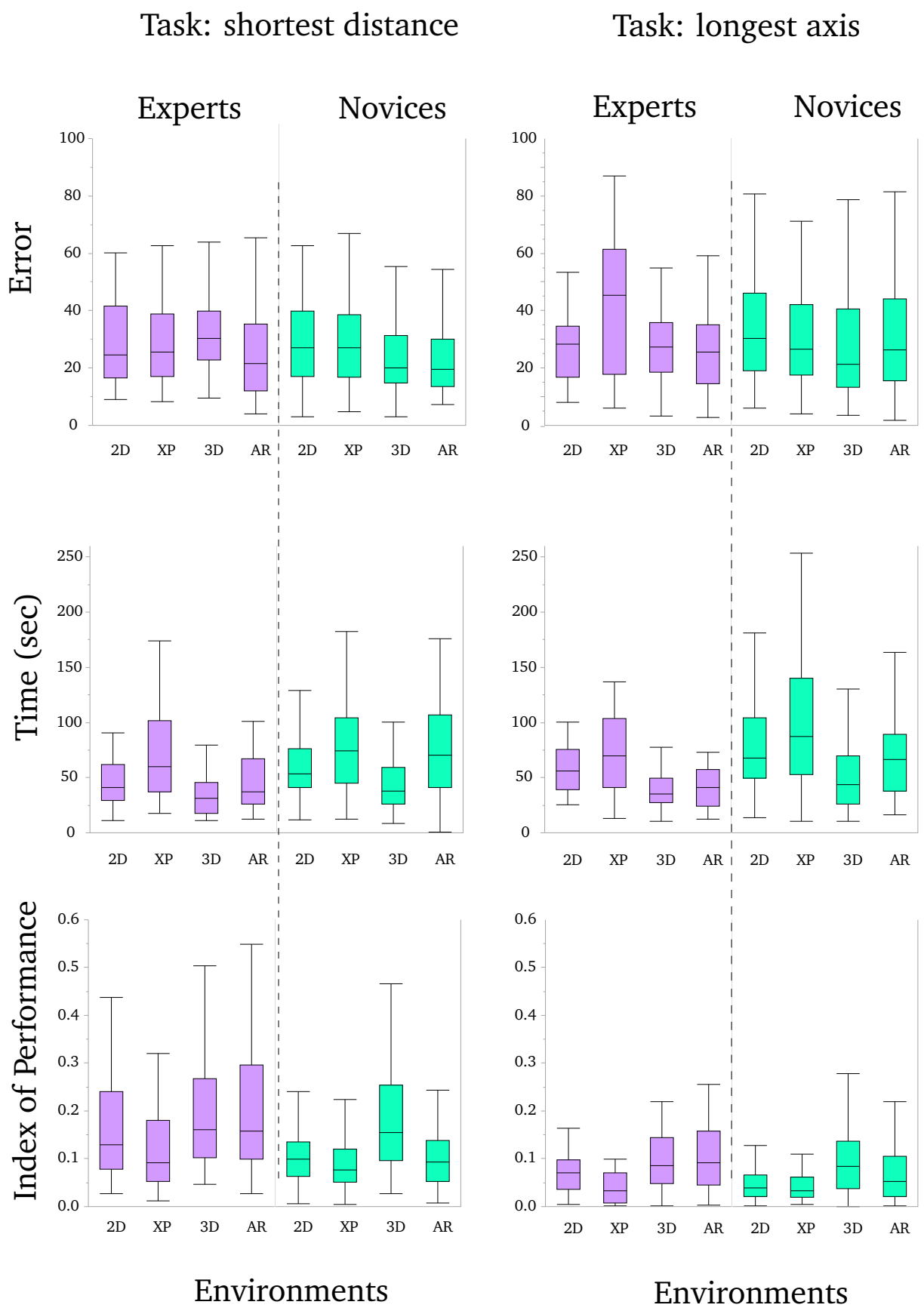


Figure 4.14: *Top*: Translational (*left)* and rotational error (*right*); *Middle*: Time for SD (*left)* and LA (*right*); *Bottom*: I_p for SD (*left)* and LA (*right*)

Table 4.2: Average rotational and translational error (phase 3)

Rotational error (^{o}), $\mu \pm \sigma$			
AR	3*D*	2*D*	*XP*
38.0 ± 24.2	41.2 ± 23.7	42.8 ± 23.2	45.0 ± 23.7

Translational error (mm), $\mu \pm \sigma$			
AR	3*D*	2*D*	*XP*
37.3 ± 31.5	39.0 ± 27.7	39.5 ± 27.0	40.6 ± 32.9

ing that users' performance in selecting the longest axis within different environments depended in part on their level of expertise. Post-hoc analysis using Tukey HSD test revealed that, regardless of the task performed, experts performed significantly faster and better (i.e. higher I_p) than novices within the AR and 2D environments (p<0.05). This improvement of I_p within the 2D environment was not surprising given the fact that experts are highly accustomed to interpreting 2D images. Furthermore, regardless of the task performed, experts performed significantly faster within the AR and 3D environments compared to the XP, whereas novices performed significantly faster within the 3D environment compared to the XP and 2D (p<0.05). In addition, for the restricted range of eccentricities displayed, no correlation was observed between the eccentricity of the tumours and performance in any of the environments.

The translational and rotational errors reported in this study appear to be relatively high for IGI applications. However, this is due to the nature of the metrics defined. For instance, one can imagine that selecting an entry point that corresponds to the second shortest distance to the target could be very distant from the optimal point of entry, resulting in a large translational error. To this aim, we defined two other metrics as the followings: a translational error as the difference (in *mm*) between the length of the optimal path and the path chosen by the user, and a rotational error as the difference (in *mm*) between the section of the optimal and chosen path cutting through the tumour. Analyzing data based on these two metrics led to very similar results, illustrating the

robustness of our initial metrics.

Overall, following observations can be made based on the results:

- *Phase* 1: the improvement of performance in 3D and AR indicates that appropriate visualization methods (i.e. AR, 3D) can significantly facilitate the perception and mental transformation of the target location and orientation and its spatial relationship with surrounding anatomical context;

- *Phase* 2: the improvement of performance in AR demonstrates that AR can significantly facilitate the process of perceptual integration and coordination between frames. In other words, AR is superior to the other planning environments in assisting novice users to generate a mental representation of the brain and transform it from a display to the patient frame of reference. In addition, observing an interaction between the task and the environment in 2D and XP indicates their lack of generic usability. In contrast, this suggests the use of AR for applications that involve tasks with multiple and possibly conflicting criteria;

- *Phase* 3: the improvement of speed in AR and 3D compared to the XP and 2D regardless of expertise and the task performed illustrates the potential impact of such environments in increasing the efficiency of planning. One unexpected result, however, was the relationship between the 2D and XP. Our conceptual models predicted that XP would have improved performance compared to 2D afforded by presenting information in a three-dimensional context. Although the method of interaction in XP was more expressive, it was also more complex, leading to lengthened response times across all participant groups. This could be mitigated in practice as users become more familiar with this mode of interactive viewing.

4.4 Conclusion

In the realm of minimally-invasive surgery, virtual environments are increasingly being developed and deployed to teach a variety of skills, from decision making to surgical dexterity. However, despite the growing body of evidence [62, 63, 64], little attention has been given to the use of virtual- and augmented-reality environments to improve general spatial abilities and skills. To this end, we designed, developed, and evaluated an augmented-reality environment by which novice physicians can practice and improve their basic spatial skills, increasing their cognitive ability to perform neurosurgical planning. Design of the proposed environment was accomplished by careful consideration of the cognitive and perceptual capacities and limitations of human observers. To evaluate our system, a number of experiments were conducted in which subjects performed relevant spatial judgment tasks using the proposed system along with conventional approaches. Our results indicate that AR environments could facilitate the task of planning brain tumour resections according to clinically-relevant criteria. Non-intuitive environments, on the other hand, were shown to be inefficient and less effective. The proposed work is a preliminary step towards the clinical use of augmented-reality environments, particularly in facilitating brain tumour resections.

4.5 Limitations and Future Work

Poor ergonomics of the AR goggles was repetitively reported to be a performance-impeding factor in the AR environment. In spite of placing multiple markers all around the phantom and goggles to increase their visibility, the issue of line-of-sight was also occasionally disrupting the experiments. In terms of experimental design, the task completion time was defined as the time taken to explore the data, interact with the phantom, and specify the response using the stylus. Therefore, computing the I_P based on the overall task completion time, though provides a good metric of performance,

limits our apprehension of the underlying sub-tasks.

In case of patient-specific data, defining a gold standard trajectory for the avoidance task was cumbersome due to the fact that multiple angles and locations could be chosen as the optimal trajectories and points of entry. For the future studies, this can be resolved by generating a consensus over the optimal trajectory and perhaps defining a gold standard *region* as opposed to a straight surgical path. In tumour resection interventions, estimating the location and size of craniotomy is another important aspect of planning. In this work, however, only the former was investigated and therefore, future studies may involve exploring the role of different planning environments in estimating the size of craniotomy. These planning environments can also be extended to include *augmented-virtuality* environments in which the patient data along with virtual representation of the phantom and stylus are shown on a computer display (a concept similar to [65] in which pre- and intra-operative data are integrated with surgical instrument tracking as an augmented-virtuality platform for mitral valve repair).

In addition to human factors, the overall performance in the proposed AR environment is the result of the system's inherent accuracy (i.e., the registration and tracking accuracy) as well as the lag between one's proprioception and the visual feedback provided through the goggles. Although these factors were not investigated for purposes of scope, they need to be taken into consideration for future studies.

Other factors need to be considered for future studies include increasing the sample size of expert population, providing training for subjects, including subjective analysis, and investigating the possible correlation between mental rotation abilities and performance.

Bibliography

[1] Public Health Agency of Canada Canadian Cancer Society. Canadian cancer statistics, 2013.

[2] American Cancer Society. Cancer facts and figures, 2013.

[3] Central Brain Tumor Registry of the United States. Primary brain and central nervous system tumors diagnosed in the united states in 2004-2008, 2012.

[4] A. Quinones-Hinojosa. *Schmidek and Sweet: Operative Neurosurgical Techniques: Indications, Methods and Results (Expert Consult - Online and Print)*. Elsevier Health Sciences, 2012.

[5] Jeffrey M Zacks and Pascale Michelon. Transformations of visuospatial images. *Behavioral and Cognitive Neuroscience Reviews*, 4(2):96–118, 2005.

[6] Christopher D Wickens, Michelle Vincow, and Michelle Yeh. Design applications of visual spatial thinking: The importance of frame of reference. *The Cambridge handbook of visuospatial thinking*, pages 383–425, 2005.

[7] Daniel Reisberg. *The Oxford Handbook of Cognitive Psychology*. Oxford University Press, 2013.

[8] Terry M Peters. Image-guidance for surgical procedures. *Physics in medicine and biology*, 51(14):R505, 2006.

[9] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[10] H.M. L. Sensorama simulator, August 28 1962. US Patent 3,050,870.

[11] Jeffrey H Shuhaiber. Augmented reality in surgery. *Archives of surgery*, 139(2):170–174, 2004.

[12] Ronald T Azuma et al. A survey of augmented reality. *Presence*, 6(4):355–385, 1997.

[13] Tobias Sielhorst, Marco Feuerstein, and Nassir Navab. Advanced medical displays: A literature review of augmented reality. *Display Technology, Journal of*, 4(4):451–467, 2008.

[14] Marta Kersten-Oertel, Pierre Jannin, and D Louis Collins. Dvv: a taxonomy for mixed reality visualization in image guided surgery. *Visualization and Computer Graphics, IEEE Transactions on*, 18(2):332–352, 2012.

[15] Cristian A Linte, Katherine P Davenport, Kevin Cleary, Craig Peters, Kirby G Vosburgh, Nassir Navab, Philip Edwards, Pierre Jannin, Terry M Peters, David R Holmes III, et al. On mixed reality environments for minimally invasive therapy guidance: Systems architecture, successes and challenges in their implementation from laboratory to clinic. *Computerized Medical Imaging and Graphics*, 37(2):83–97, 2013.

[16] Timothy Poston and Luis Serra. The virtual workbench: dextrous vr. In *Proceedings of the conference on Virtual reality software and technology*, pages 111–121. World Scientific Publishing Co., Inc., 1994.

[17] Axel Thomas Stadie, Ralf Alfons Kockro, Robert Reisch, Andrei Tropine, Stephan Boor, Peter Stoeter, and Axel Perneczky. Virtual reality system for planning minimally invasive neurosurgery. 2008.

[18] Cristian Luciano, Pat Banerjee, Lucian Florea, and Greg Dawe. Design of the immersivetouch: a high-performance haptic augmented virtual reality system. In *Proceedings of the 11th international conference on human-computer interaction, Las Vegas, Nevada*, 2005.

[19] Gabor Fichtinger, Anton Deguet, Ken Masamune, Emese Balogh, Gregory S Fischer, Herve Mathieu, Russell H Taylor, S James Zinreich, and Laura M Fayad. Image overlay guidance for needle insertion in CT scanner. *Biomedical Engineering, IEEE Transactions on*, 52(8):1415–1424, 2005.

[20] Ken Hinckley, Randy Pausch, J Hunter Downs, Dennis Proffitt, and Neal F Kassell. The props-based interface for neurosurgical visualization. *Studies in health technology and informatics*, 39:552–562, 1997.

[21] WEL Grimson, RJFA Kikinis, Ferenc A Jolesz, and PM Black. Image-guided surgery. *Scientific American*, 280(6):54–61, 1999.

[22] Ralf A Kockro, Yeo Tseng Tsai, Ivan Ng, Peter Hwang, Chuangui Zhu, Kusuma Agusanto, Liang Xiao Hong, and Luis Serra. Dex-ray: Augmented reality neurosurgical navigation with a handheld video probe. *Neurosurgery*, 65(4):795–808, 2009.

[23] Philip J Edwards, Andrew P King, Calvin R Maurer Jr, Darryl A De Cunha, David J Hawkes, Derek LG Hill, Ronald P Gaston, Michael R Fenlon, A Jusczyzck, Anthony J Strong, et al. Design and evaluation of a system for microscope-assisted guided interventions (MAGI). *Medical Imaging, IEEE Transactions on*, 19(11):1082–1093, 2000.

[24] Ramin Shahidi, Michael R Bax, Calvin R Maurer Jr, Jeremy A Johnson, Eric P Wilkinson, Bai Wang, Jay B West, Martin J Citardi, Kim H Manwaring, and Rasool Khadem. Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. *Medical Imaging, IEEE Transactions on*, 21(12):1524–1535, 2002.

[25] Christoph Bichlmeier, Sandro Michael Heining, Marco Feuerstein, and Nassir Navab. The virtual mirror: a new interaction paradigm for augmented reality environments. *Medical Imaging, IEEE Transactions on*, 28(9):1498–1510, 2009.

[26] Calvin R Maurer Jr, Frank Sauer, Bo Hu, Benedicte Bascle, Bernhard Geiger, Fabian Wenzel, Filippo Recchi, Torsten Rohlfing, Christopher R Brown, Robert J Bakos, et al. Augmented-reality visualization of brain structures with stereo and kinetic depth cues: system description and initial evaluation with head phantom. In *Medical Imaging 2001*, pages 445–456. International Society for Optics and Photonics, 2001.

[27] Mirna Lerotic, Adrian J Chung, George Mylonas, and Guang-Zhong Yang. Pq-space based non-photorealistic rendering for augmented reality. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 102–109. Springer, 2007.

[28] Perrine Paul, Oliver Fleig, and Pierre Jannin. Augmented virtuality based on stereoscopic reconstruction in multimodal image-guided neurosurgery: Methods and performance evaluation. *Medical Imaging, IEEE Transactions on*, 24(11):1500–1511, 2005.

[29] Damini Dey, David G Gobbi, Piotr J Slomka, Kathleen JM Surry, and Terry M. Peters. Automatic fusion of freehand endoscopic brain images to three-dimensional

surfaces: creating stereoscopic panoramas. *Medical Imaging, IEEE Transactions on*, 21(1):23–30, 2002.

[30] Philip J Edwards, Laura G Johnson, David J Hawkes, Michael R Fenlon, Anthony J Strong, and Michael J Gleeson. Clinical experience and perception in stereo augmented reality surgical navigation. In *Medical Imaging and Augmented Reality*, pages 369–376. Springer, 2004.

[31] Laura G Johnson, Philip Edwards, and David Hawkes. Surface transparency makes stereo overlays unpredictable: The implications for augmented reality. *Studies in health technology and informatics*, pages 131–136, 2003.

[32] Mary Hegarty, Madeleine Keehner, Cheryl Cohen, Daniel R Montello, and Yvonne Lippa. The role of spatial cognition in medicine: Applications for selecting and training professionals. *Applied spatial cognition*, pages 285–315, 2007.

[33] Jane M Connor and Lisa A Serbin. Visual-spatial skill: Is it important for mathematics? can it be taught. *Women and mathematics: Balancing the equation*, pages 151–174, 1985.

[34] David F Lohman. Spatial abilities as traits, processes, and knowledge. 1988.

[35] Maryann Baenninger and Nora Newcombe. The role of experience in spatial test performance: A meta-analysis. *Sex roles*, 20(5-6):327–344, 1989.

[36] Ross Vasta, Jill A Knott, and Christine E Gaze. Can spatial training erase the gender differences on the water-level task? *Psychology of Women Quarterly*, 20(4):549–567, 1996.

[37] Roy Eagleson. Perceptual capacitites and constraints in augmented reality biomedical displays. In *Innovations in Patient Care: Engineering and Physical Sci-*

ences in Medicine. Australian Biomedical Engineering Conference. Christchurch, N.Z., 2008.

[38] Elvis CS Chen, Kripasindhu Sarkar, John SH Baxter, John Moore, Chris Wedlake, and Terry M Peters. An augmented reality platform for planning of minimally invasive cardiac surgeries. In *SPIE Medical Imaging*, pages 831617–831617. International Society for Optics and Photonics, 2012.

[39] Rudolph E Kalman and Richard S Bucy. New results in linear filtering and prediction theory. *Journal of basic engineering*, 83(1):95–108, 1961.

[40] John SH Baxter, Terry M Peters, and Elvis CS Chen. A unified framework for voxel classification and triangulation. In *SPIE Medical Imaging*, pages 796436–796436. International Society for Optics and Photonics, 2011.

[41] Colin Ware. *Information visualization: perception for design*. Elsevier, 2012.

[42] Ernst Kruijff, J Edward Swan II, and Steven Feiner. Perceptual issues in augmented reality revisited. In *ISMAR*, volume 9, pages 3–12, 2010.

[43] David Drascic and Paul Milgram. Perceptual issues in augmented reality. In *Electronic Imaging: Science & Technology*, pages 123–134. International Society for Optics and Photonics, 1996.

[44] Mark Mon-Williams and James R Tresilian. Ordinal depth information from accommodation? *Ergonomics*, 43(3):391–404, 2000.

[45] Shojiro Nagata. How to reinforce perception of depth in single two-dimensional pictures. *Proceedings of the society for information display*, 25(3):239–246, 1984.

[46] Kamyar Abhari, John SH Baxter, Roy Eagleson, Terry Peters, and Sandrine de Ribaupierre. Perceptual enhancement of arteriovenous malformation in MRI angiog-

raphy displays. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8318, page 8, 2012.

[47] Kamyar Abhari, John SH Baxter, Ali R Khan, Terry Peters, Sandrine de Ribaupierre, and Roy Eagleson. Visual enhancement of mr angiography images to facilitate planning of arteriovenous malformation interventions. *Submitted to ACM Transactions on Applied Perception*, 2014.

[48] Oliver Kutter, André Aichert, Christoph Bichlmeier, Jörg Traub, Sandro Michael, Ben Ockert Heining, Ekkehard Euler, and Nassir Navab. Real-time volume rendering for high quality visualization in augmented reality. 2008.

[49] Brian Rogers, Maureen Graham, et al. Motion parallax as an independent cue for depth perception. *Perception*, 8(2):125–134, 1979.

[50] Hiroshi Ono and Martin J Steinbach. Monocular stereopsis with and without head movement. *Perception and Psychophysics*, 48(2):179–187, 1990.

[51] V Cornilleau-Peres and J Droulez. The visual perception of three-dimensional shape from self-motion and object-motion. *Vision research*, 34(18):2331–2336, 1994.

[52] Mark R Mine, Frederick P Brooks Jr, and Carlo H Sequin. Moving objects in space: exploiting proprioception in virtual-environment interaction. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 19–26. ACM Press/Addison-Wesley Publishing Co., 1997.

[53] Leonard J Williams. Tunnel vision induced by a foveal load manipulation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 27(2):221–227, 1985.

[54] Michael Bajura, Henry Fuchs, and Ryutarou Ohbuchi. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. In *ACM SIGGRAPH Computer Graphics*, volume 26, pages 203–210. ACM, 1992.

[55] Christoph Bichlmeier, F Wimme, Sandro Michael Heining, and Nassir Navab. Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, pages 129–138. IEEE, 2007.

[56] Steve Mann and Steve Mann Nnlf. Mediated reality. 1994.

[57] Ruth Rosenholtz, Yuanzhen Li, and Lisa Nakano. Measuring visual clutter. *Journal of Vision*, 7(2):17, 2007.

[58] Kamyar Abhari, John SH Baxter, Elvis S Chen, Ali R Khan, Chris Wedlake, Terry M Peters, Sandrine de Ribaupierre, and Roy Eagleson. Use of a mixed-reality system to improve the planning of brain tumour resections: Preliminary results. In *Augmented Environments for Computer-Assisted Interventions*, pages 55–66. Springer, 2013.

[59] Kamyar Abhari, John SH Baxter, Elvis S Chen, Ali R Khan, Chris Wedlake, Terry Peters, Roy Eagleson, and Sandrine de Ribaupierre. The role of augmented reality in training the planning of brain tumor resection. In *Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*, pages 241–248. Springer, 2013.

[60] Paul M Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6):381, 1954.

[61] Roy Eagleson. Surgical task decomposition for simulator design and evaluation. In *Lecture Series on Effective Systems for Patient Specific Surgical Simulation*. CSTAR Multimedia Theater, University Hospital, London, Canada, 2013.

[62] Andreas Dünser, Karin Steinbügl, Hannes Kaufmann, and Judith Glück. Virtual and augmented reality as spatial ability training tools. In *Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI*, pages 125–132. ACM, 2006.

[63] Hannes Kaufmann and Dieter Schmalstieg. Mathematics and geometry education with collaborative augmented reality. *Computers & Graphics*, 27(3):339–345, 2003.

[64] David Waller, Earl Hunt, and David Knapp. The transfer of spatial knowledge in virtual environment training. *Presence: Teleoperators and Virtual Environments*, 7(2):129–143, 1998.

[65] John T Moore, Michael WA Chu, Bob Kiaii, Daniel Bainbridge, Gerard Guiraudon, Chris Wedlake, Maria Currie, Martin Rajchl, Rajni V Patel, and Terry M Peters. A navigation platform for guidance of beating heart transapical mitral valve repair. *Biomedical Engineering, IEEE Transactions on*, 60(4):1034–1040, 2013.

Chapter 5

A Complete Simulation Environment for Vertebroplasty Procedure

This chapter is a result of collaboration with *Navigated Augmented Reality Visualization Systems* (NARVIS) *laboratory* at Technical University of Munich (TUM), Germany and is adapted from '*Vertebroplasty Performance on Simulator for* 19 *Surgeons Using Hierarchical Task Analysis*'[1].

My contribution to this chapter involved assisting in (i) designing and conducting experiments, (ii) analyzing data (objective measures), and (iii) writing manuscripts.

5.1 Introduction

5.1.1 Clinical Motivation: Vertebroplasty

Osteoporosis is a disease characterized by low bone density and mass, leading to deterioration of bone tissue, increased bone fragility, and risk of fracture [1]. Osteoprosis

[1]Abhari. K.*, Wucherer P.*, Stefan P.*, Fallavollita P., Weigl M., Lazarovici M., Winkler A., Weidert S., Eagleson R., de Ribaupierre S., Peters T., Navab N., "Vertebroplasty Performance on Simulator for 19 Surgeons Using Hierarchical Task Analysis", Submitted to *IEEE Transaction on Medical Imaging*, (2014) **equal contribution*

affects 12% of population within the age range of $50 \sim 79$ in Europe, and approximately 1.4 million Canadians [2] and 10 million Americans [3]. These patients have 23% higher rate of mortality compared to corresponding normal population [4, 5], imposing a social and an economic concern. Percutaneous *Vertebroplasty* is an outpatient minimally invasive procedure to treat vertebral fractures caused by osteoprosis. Originally developed in 1987 to treat destruction of vertebrae caused by haemangiomas, today a wide range of vertebral fractures can be treated by vertebroplasty [6]. In this procedure, a trocar is guided under C-arm (or sometimes O-arm) fluoroscopic guidance into the collapsed vertebral body [7]. After reaching the point of interest, a catheter is inserted in order to inject polymethylmethacrylate (PMMA) bone cement into the target vertebra. PMMA quickly dries after and forms a support structure to provide stabilization. Kyphoplasty is an extension of the vertebroplasty technique involving the same interventional approach but using a balloon catheter to first expand the vertebral body before injection of PMMA. Although commonly performed, the complication rate of vertebroplasty is remarkably high [7]. Cement leakage due to needle misplacement may result in serious side effects such as infection, bleeding, numbness, headache, and paralysis [8] (occurs in $>$2% in patients with fractures due to cancer [9]). Appropriate training is necessary to reduce and eventually eliminate such intra- and postoperative complications.

5.1.2 Background

Simulation in Surgical Training

Surgical training is predominantly accomplished using the *Halstedian* apprenticeship model [10], in which real operative scenarios provide the context for learning. Vertebroplasty procedure is no exception, where trainees acquire necessary skills by observing experts and then gradually gain hands-on experience by practicing on cadavers and

patients [11]. Unfortunately, in such a training model, controlling the curriculum delivery is impossible, and in extreme cases, patient safety might be compromised. This realization has motivated alternative approaches that allow trainees to acquire experience before being exposed to real surgical scenarios. Among available techniques, surgical simulators have the potential to revolutionize the surgical education by providing self-paced learning environments to easily examine and practice different surgical approaches. Recent studies have shown simulation training can reduce the initial training errors made on patients [12, 13], and thus, they could (and should) play an expanding role in medical education. Training using simulators, if appropriately utilized and integrated into the educational curriculum, is shown to improve the surgical outcome, and most importantly, improve patient safety [14].

Non-Technical Surgical Skills

Traditionally, surgical education has focused on developing knowledge and technical skills. As a result, surgical simulators are often designed to improve dexterity, motor skills, and clinical expertise. In addition to technical skills, there is a growing body of evidence supporting the impact of *non-technical skills* on patient safety, such as decision making, situation awareness, and coping with pressure. In fact, many adverse events in the operation rooms are directly caused by non-technical aspects of surgery, such as miscommunication [15], perceptual [16], and cognitive errors [17], rather than a lack of technical expertise. According to a study performed by Gawande *et al.*, out of 100 surgical error incidents, 16% of cases were reported to be caused by interruptions and distractions [15] . In their work, it was noted that surgeons appear to be “particularly prone to underestimating the influence of factors such as fatigue and interruptions”. Cognitively speaking, these factors can manifest themselves as *extraneous loads*, degrading the performance (refer to 1.2.2 for details). It is well known, however, via enough training under appropriate circumstances, trainees can organize

information into *schemas* - or chunks of knowledge - to comprehend and conceptualize [18]. In other words, it is necessary to train residents to acquire not only technical but also non-technical skills in order to properly cope with crises and interruptions in the OR. Although this might have been acknowledged in the literature, systematic studies are scarce, prohibiting its wide acceptance by the community. This also holds true for vertebroplasty procedures in which *multitasking* is essential for navigating through the patient body. Wilson *et. al* [19] have shown that multitasking in the OR significantly increases the subjective experience of distraction and consequently increases the cognitive load, causing surgical performance to break down. Thus, extraneous interruptions and distractions in vertebroplasty can seriously disturb the surgical outcome at every stage of navigation, targeting, and imaging.

5.1.3 Related Work

Simulation in Health Care

Broadly speaking, simulation training in health care can be classified into three domains [20]: (i) Standardized patients and computerized mannequins to teach basic clinical skills such as cardiovascular examination, resuscitation training, or assessment of anaesthetists (e.g. [21, 22, 23, 24]); (ii) simulated environments to teach and assess surgical skills, ranging from animal and cadaver tissue models to computerized simulators (e.g. [25, 26, 27, 28, 29]); and (iii) simulation environments (mainly mannequin-based) to teach team-based and crisis management skills in complex scenarios (e.g. [30, 31, 20, 32]). Many agree, however, that while computerized simulators such as *virtual environments* offer a cost-effective and efficient alternative to traditional training methods, *mannequin-based* simulators can facilitate the training of non-technical skills (e.g. [32]), such as coping with interruptions. While these two classes of simulators have been extensively studied, little research has been devoted to development

and evaluation of systems that encapsulate both techniques.

Vertebroplasty Simulator

Virtual simulators for vertebroplasty or similar minimally-invasive spinal procedures are surprisingly rare. In orthopedics, virtual simulators with tactile feedback have been developed for craniofacial surgery [33], hip fracture surgery [34], spine fusion [35], and insertion of pedicle screws [36]. Chui *et al.* have developed a vertebroplasty simulator with the fluoroscopic guidance, incorporating tactile feedback using a haptic glove [25]. These simulators are often designed as a standalone training environment for single users, focusing on improving underlying surgical skills.

The simulation environment used in this study consists of a mannequin in conjunction with an augmented-virtuality (AV) environment, specifically designed for training vertebroplasty procedures, which has been previously developed and validated in [37] and [38]. The simulator is equipped with haptic feedback, displaying (preoperative) X-ray images with the trocar superimposed in real-time (Figure 5.1). In [37], four expert surgeons were asked to perform vertebroplasty using a simulator while their behaviour and workflow responses were investigated under certain adverse events such as cement leakage. The qualitative feedback from the participants supported the content validity of the system, postulating the necessity of using mannequin, visual, and haptic feedback within an immersive environment for great degree of realism.

5.1.4 Rationale, Hypothesis, and Objectives

We believe that non-technical skills, such as team training and coping with interruptions, are essential for training novice surgeons and so need to be incorporated into surgical training curriculum. To this end, the objective of this work is to assess novice surgeons' performance and cognitive load under unanticipated events, such as interruptions, while placed within a fully immersive environment. We hypothesize that

Figure 5.1: X-ray images of the trocar and spine anatomy

such extraneous distractions increase the cognitive load, degrading performance in vertebroplasty procedures. To support this, both subjective and objective analyses of performance are conducted.

5.1.5 Contributions

In this work, we present a unique approach to asses the influence of participants' non-technical (cognitive) skills on their performance while conducting a vertebroplasty procedure within a complete simulation environment. The simulation setup involves a mannequin-based AV environment with physiological responses, addressing a broad spectrum of human sensory input, including visual, tactile, and auditory channels. The proposed simulation environment is not explicitly designed for vertebroplasty, and so can be extended to similar scenarios. Nevertheless, this work could potentially make a strong case for training non-technical skills, leading to stronger foundation for credentialing, and eventually reducing the occurrence of adverse events in the OR.

5.2 Methods and Materials

An evaluation study was conducted with 19 orthopedic residents in order to:

i provide a qualitative measure of usability;

ii assess technical performance of the novice surgeons in presence of interruptions;

iii explore the relationship between mental workload and surgical performance within a fully-immersed environment.

5.2.1 Simulator Components

The simulator employed in this study had developed by our collaborators at NARVIS laboratory (TÜM, Munich, Germany) with the tactile feedback provided by an off-the-shelf haptic device (Novint Falcon, Novint Technologies Inc., NY, USA). Volume Haptics Toolkit (H3D API, SenseGraphics, Stockholm, Sweden) was used to load, handle, and process data, producing visual and haptic feedback.

For the purpose of this study, the trocar was attached to a haptic end-effector generated using rapid prototyping technology, and inserted through a pad positioned on the mannequin, simulating a vertebroplasty procedure (Figures 5.2 and 5.3). The translational force feedback (3DOF) was generated along the axis specified by the angle of entry. Advancing the trocar would generate a tactile feedback with a sensation corresponding to the tissue type (e.g., bone, fat, air). Accordingly, this process would provide a stiff sensation when the trocar goes through the cortical bone structure, with a significant drop in resistance after penetrating through the cancellous tissue. The visualization interface involved rendering the projection of the trocar as well as the patient-specific X-ray images of the spine anatomy (Figure 5.1). Using a foot-pedal, participating surgeons could acquire anterior-posterior (AP) or lateral (LAT) view of X-ray images, displayed on a monitor placed on the opposite side of the patient. In ad-

dition, a virtual C-arm fluoroscope was also integrated into the visualization platform whose location and orientation can be adjusted using a dedicated control box (Figure 5.4). For the technical details, refer to [37] and [38].

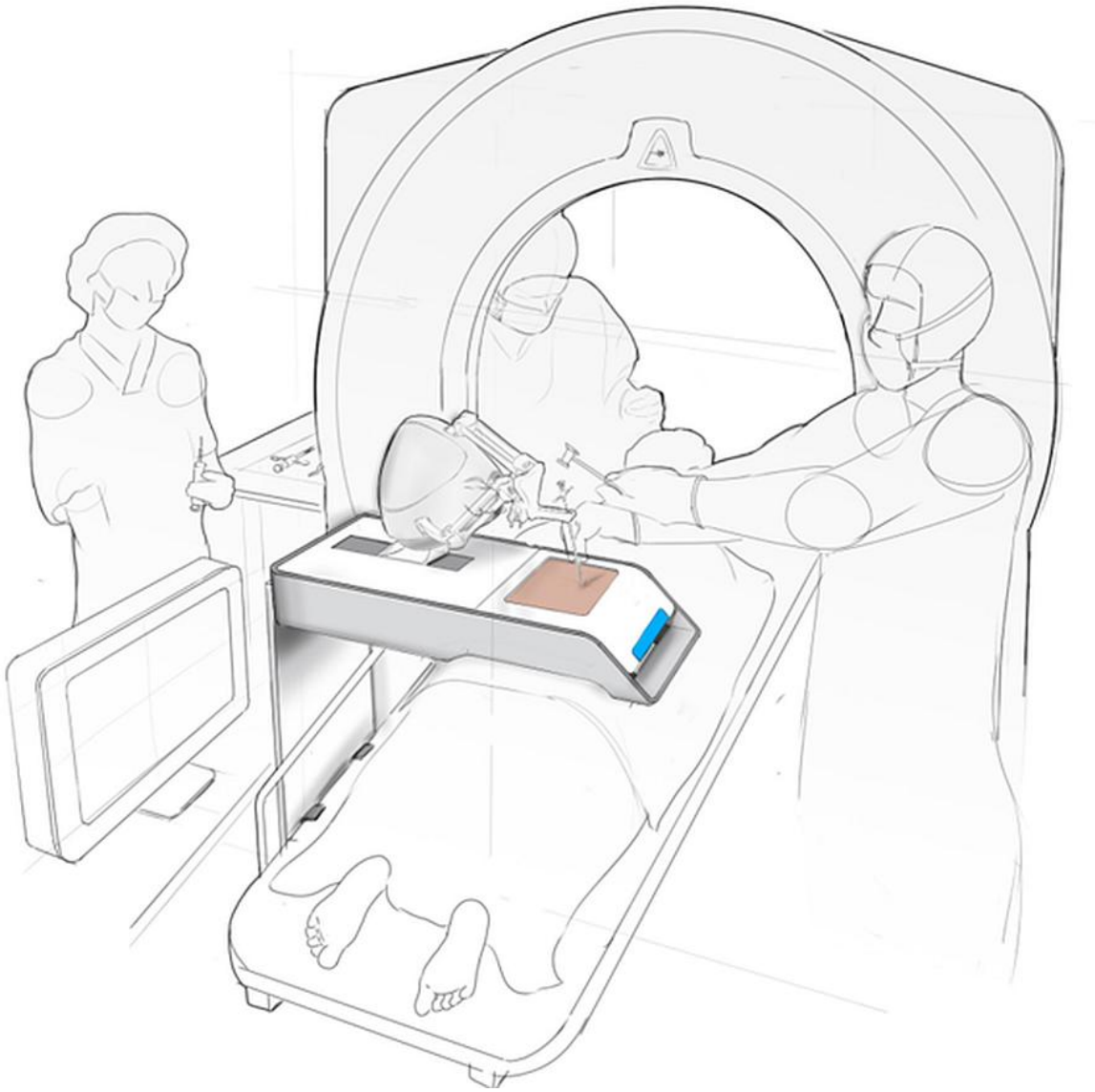


Figure 5.2: The simulation environment for the vertebroplasty procedures

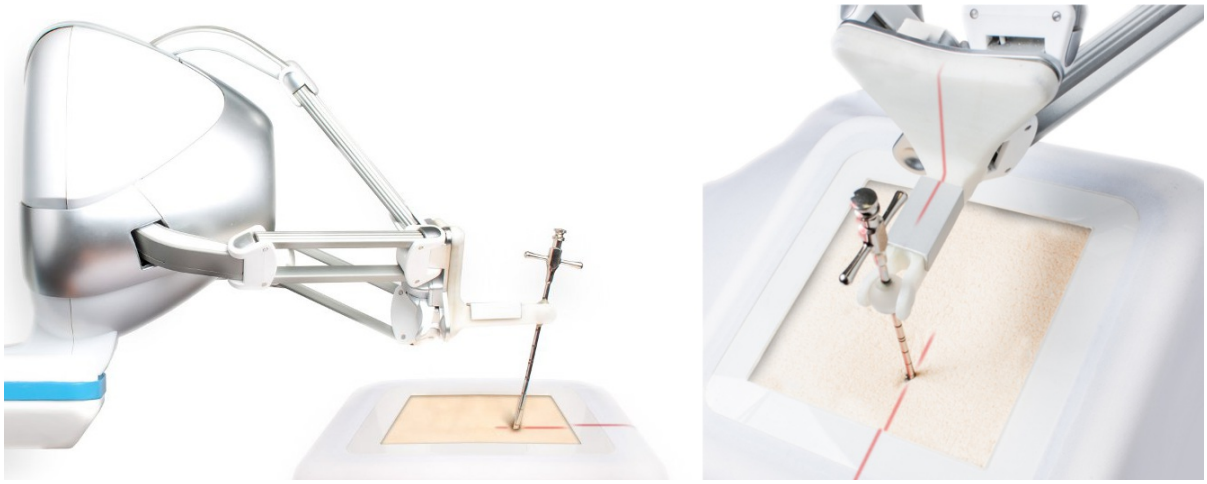


Figure 5.3: The trocar is attached to a haptic end-effector before being inserted through a pad positioned on top of the simulated patient

5.2.2 Methodology

19 junior surgeons (12 male/7 female, with 12.5 ± 9.16 months of training in orthopedic surgery) with no prior experience in performing vertebroplasty were recruited and divided into two groups with equal distribution of male/female participants. All of the participants went through a briefing about the procedure by a senior surgeon overseeing the study. Experiments consisted of a *training* phase and an *immersion* phase. The *training* phase consisted of the following 3 independent runs:

- *Informal training*: Subjects were asked to familiarize themselves with the simulator environment;

- *Training, run 1*: Subjects were asked to use the 'gold standard' surgical path - which was shown on the X-ray images - as the guide to navigate to the target vertebra. The gold standard was identified and confirmed by two expert orthopedic surgeons with extensive experience in performing vertebroplasty;

- *Training, run 2*: Subjects were asked to navigate the trocar to the target vertebra with no guidance.

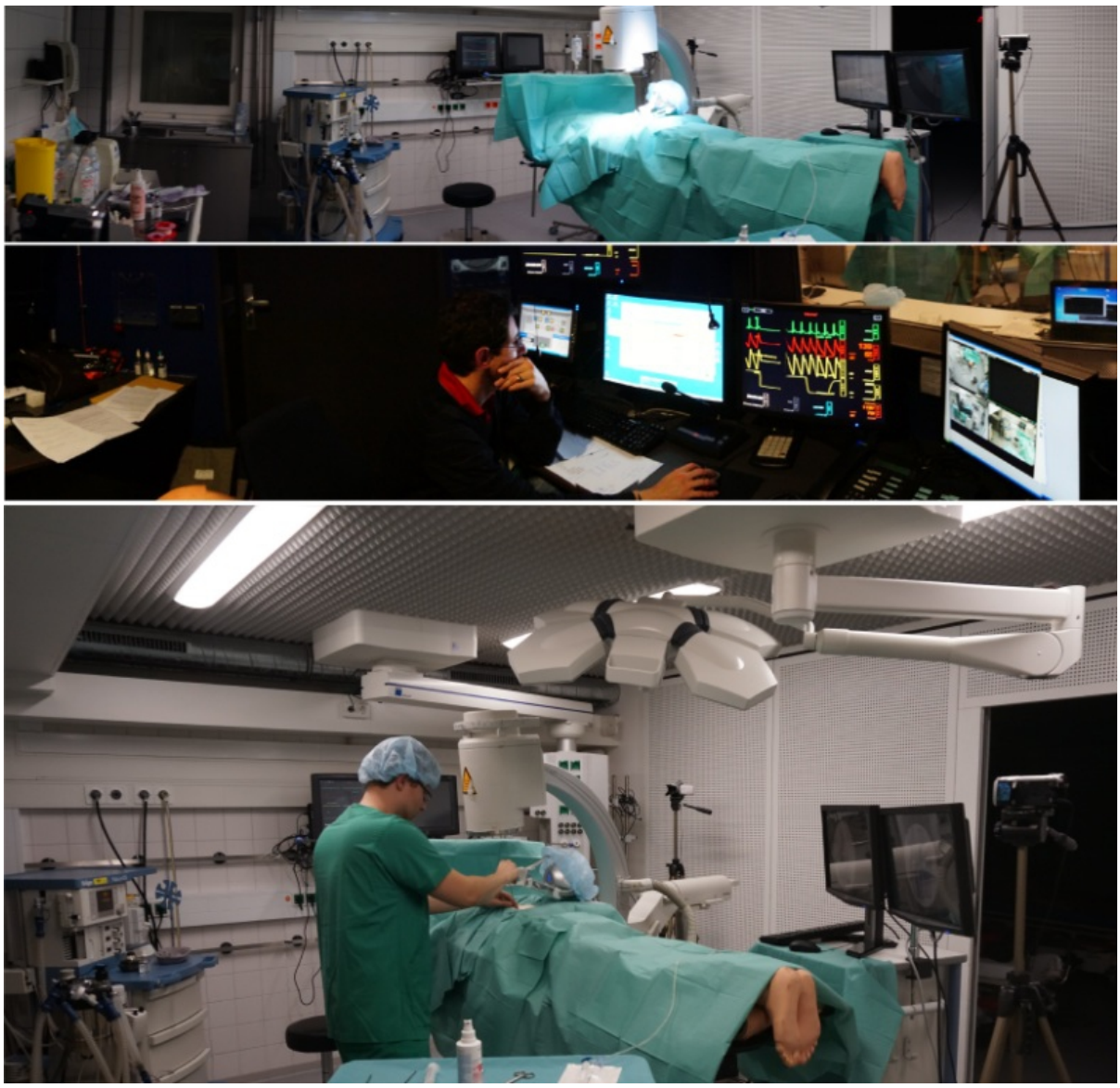


Figure 5.4: *top-bottom*: The vertebroplasty simulation environment consisting of mannequin, VR simulator, mobile C-arm fluoroscope, real medical instrument and a broad spectrum of human sensory channels such as tactile, auditory and visual in real-time. *middle*: Independent control room varying mannequin physiology and initialize crisis scenarios.

During the *immersion* phase, participants were asked to perform vertebroplasty after being fully immersed in the real simulation environment by wearing surgical gloves, a surgical mask, a lead vest, and a cover. They also underwent a short briefing about the patient (e.g. name, age, bone structure) and his status (e.g. locally anesthetized). The entry point was determined in advance by two expert surgeons and so was consistent

throughout the experiments. Depending on the group, the following *interruptions* were introduced during the procedure when the trocar reached a certain depth ($^1/_2$ of total depth) within the patient's body:

Group 1: 'phone call': The scrub nurse responded to a phone call from the head of the department who wanted to ask the participant to prepare a presentation for a meeting.

Group 2: 'patient discomfort': The patient's heart rate started to rise and the patient began to voice discomfort. The participants, therefore, had to inject more local anesthetics to relieve the pain;

The procedure resumed when the phone call ended or the patient reported no further discomfort. It should be noted that in order to avoid the effect of fatigue, subjects were instructed to take short breaks between all four phases of study.

5.2.3 Equipment and Environment

The Institute for Emergency Medicine and Management[2] provided the necessary infrastructure for conducting the evaluation study including a replication of a real operating theater with an adjacent control room, a mobile C-arm fluoroscope, and a computerized mannequin simulator (HAL S2001, Gaumard®Scientific, Miami, USA) enveloped with surgical drapes (Figure 5.4). The mannequin's physiologic parameters could be manipulated through an interface located in the control room (Figure 5.5). Two monitors were also integrated for displaying the LAT and AP view of the operating site. The X-ray image used in this study was acquired preoperatively for an actual vertebroplasty procedure. Additionally, an 'anesthesiologist' and a 'scrub nurse' were also assisted the participants, conducting vertebroplasty.

[2]Institut für Notfallmedizin und Medizinmanagement, Ludwig Maximilian University of Munich, Munich, Germany

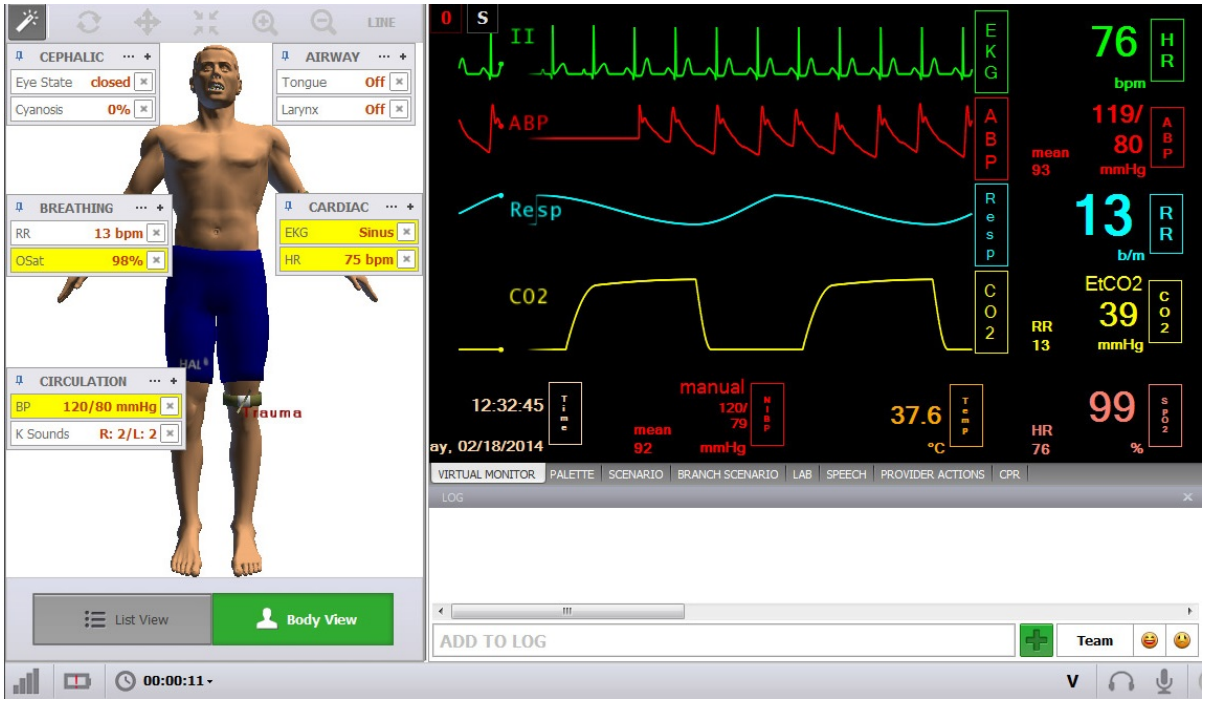


Figure 5.5: Patient's physiologic parameters could be manipulated through an interface located in the control room

5.2.4 Analysis

Objective Analysis

The position of the tip of the trocar (Figures 5.6, 5.7) and the time elapsed were recorded to compute the accuracy and speed of the participants at each stage of the experiment. The time taken to deal with the interruptions, i.e. speaking over the phone or administrating anesthesia, was excluded in our analysis. In addition, the number of X-ray shots and their duration were recorded to measure the overall administered radiation dose. Results are summarized in section 5.3.1.

Subjective Analysis

After each experiment, subjects were asked to report their subjective workload (0 :very low ~ 8 :very high) by filling out a questionnaire, similar to that described in [19] by Wilson et al.; that is a surgery-specific, multidimensional workload measure based on

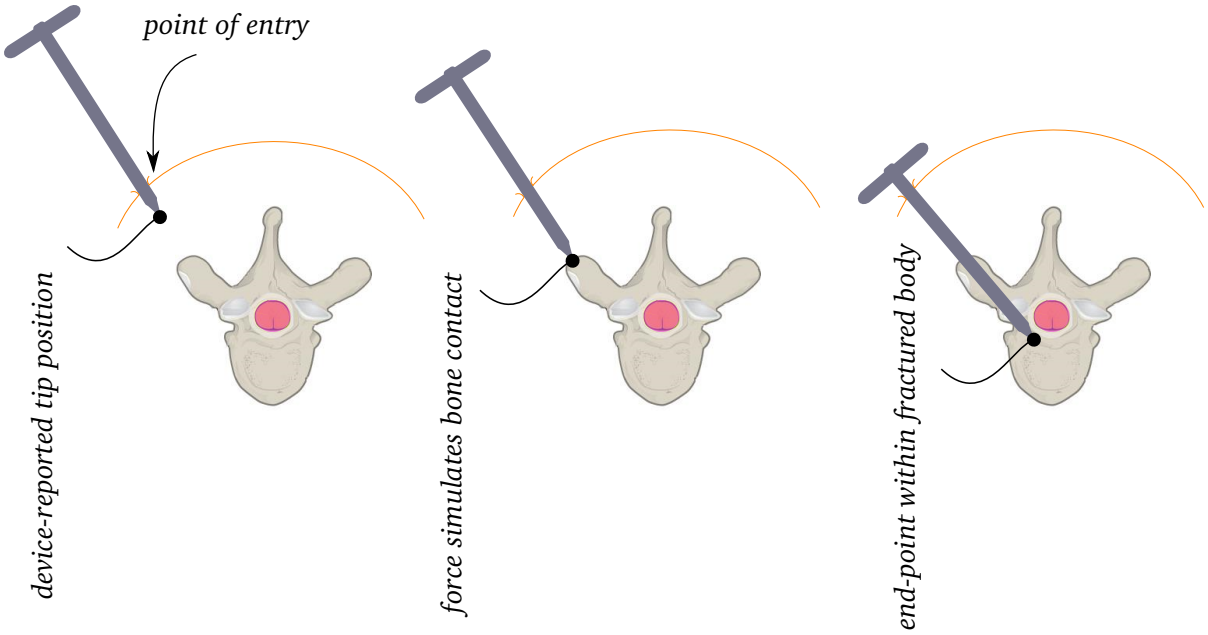


Figure 5.6: The position of the trocar throughout the procedure

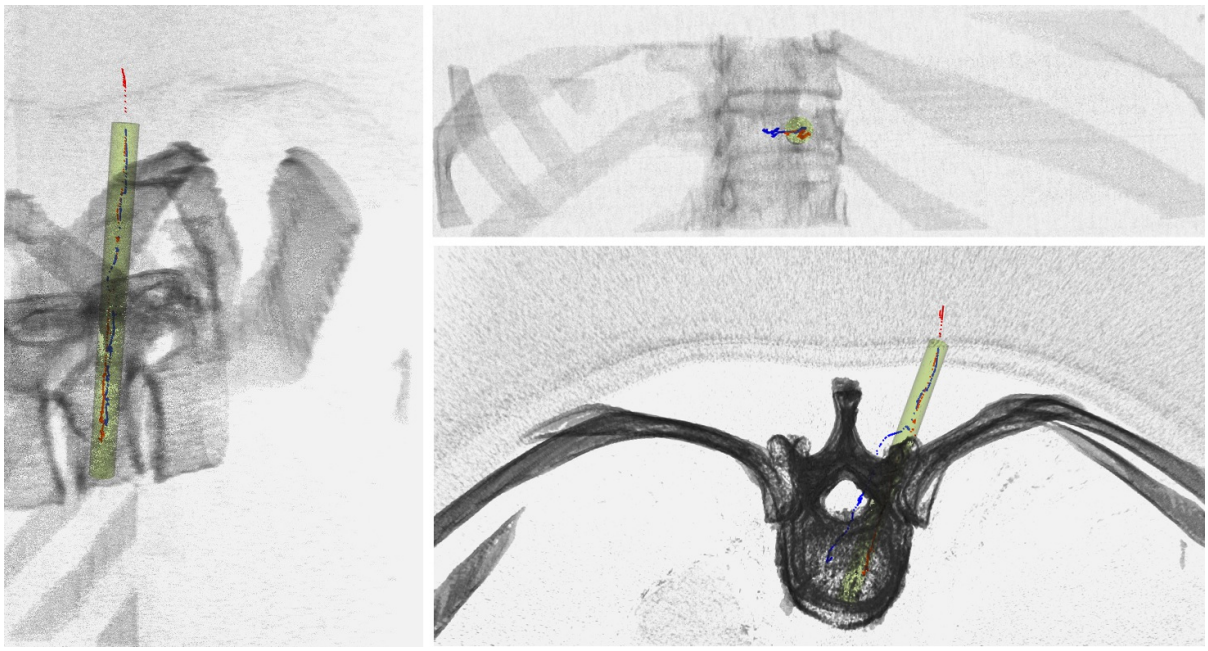


Figure 5.7: Example viewpoints of two surgeon trocar insertions versus ground truth trajectories

the NASA-TLX [39], including measures of *mental demand*, *physical demand*, *temporal demand*, *complexity*, and *situational stress* (for more details about NASA-TLX refer to section 1.2.2). Subjective responses to these items were then aggregated to compute

the overall workload. Face validity of the system - that is a subjective estimate of the extent to which the simulator seems *real* to an expert - was also evaluated via 13 questions on a 5-point Likert scale (1 :not realistic ~ 5 :very realistic). Refer to appendix B for examples of questionnaires designed by by our collaborators at NARVIS and administered in this study.

5.3 Results and Discussion

5.3.1 Results

The analysis of performance is shown in Table 5.1[3]. Note that the informal training was merely for the subjects to become familiar with the simulator and therefore its data are not included in our performance analysis.

		n	RMSD error (mm)	X-ray exposure (sec)	Time (sec)
Training run 1		19	4.3 ± 2.2	16.3 ± 14.1	126.6 ± 70.8
Training run 2		19	4.6 ± 2.8	9.9 ± 13.2	81.8 ± 59.2
Immersive	Crisis 1	8	6.7 ± 3.1	12.7 ± 13.0	305.5 ± 208.4
	Crisis 2	8	5.9 ± 2.8	11.8 ± 15.1	352.5 ± 256.4

Table 5.1: Performance Analysis

A significant difference was observed between groups in terms of task completion time (Kruskal-Wallis[4]: Chi-Square = 26.9, p<0.001). Post-hoc analysis using Tukey-HSD test reveals that the significant difference was between the two training and the immersive runs (p<0.05). Furthermore, no significant difference was observed between different phases in terms of error and X-ray exposure time. Nevertheless, the following observations can be made:

- *Time*: Although training subjects resulted in a decrease in time (training run 2

[3]Due to technical difficulties, position information were not properly recorded for 3 participants during the immersive experiments and therefore were excluded from the performance analysis.

[4]Non-parametric equivalent to one-way ANOVA

vs. training run 1), subjects performed significantly slower when placed in an immersive environment, facing with a crisis.

- *Accuracy*: After the training, subjects performance with no visual aid reached the same level of accuracy as with a visual aid. Subjects tended to be less accurate during the crises but the results were not significant.

- *X-ray*: Similar to task completion time, training resulted in a decrease and facing with crises led into an increase in the X-ray exposure time. However, the results were not significant.

The face validity of the system was reported as 4.41 ± 0.41 for the training and 4.22 ± 0.67 for the crisis simulation. The realism of the phone call and patient discomfort were respectively scored as 4.38 ± 0.55 and 4.16 ± 0.37 on a $0-5$ scale. Subjects reported significantly lower mental workload during the training sessions than the immersive runs (training: 12.94 ± 4.70, immersive:15.06 ± 6.61, paired t-test: $t = -2.53$, $p<0.02$), with no significant difference between the two crises (phone call ($N = 8$): 3.56 ± 2.26, patient discomfort ($N = 8$): 1.96 ± 0.93, $t = 1.86$, $p = 0.08$[5]). Partial correlation analysis (controlled for professional tenure) illustrates a negative correlation between subjects' mental workload and the their level of performance ($r = -0.52$, $p = .049$).

Based on these results, the following observations can be made:

- Simulating crises within an immersive environments imposes a significant increase in the mental workload.

- Increase in the mental workload negatively influence the surgical performance.

[5]Reported numbers are only one item of SURG-TLX index (perceived interruptions during the crisis)

5.3.2 Discussion

Performing a vertebroplasty procedure comprises a few major tasks: (*i*) formulating the location/angle of point of entry; (*ii*) guiding the needle into the fractures vertebra under flouroscopic guidance or *needle guidance*; (*iii*) stabilizing the vertebral bone by injecting bone cement and so filling the spaces within; and finally (*iv*) visually confirming the injection, inspecting for leaks, and removing the catheter. Despite the capability of the simulator to simulate and practice all three major tasks, we only focus on *needle guidance* for the purpose of this study, which can be broken down into the following sub-tasks (For the sake of consistency, however, the initial position/angle of the trocar was identical across subjects):

- Puncturing the skin and slowly inserting the needle;
- Guiding the needle towards the fractured vertebra;
- Controlling the view by positioning the c-arm[6] and applying X-ray;
- Inserting the needle into the vertebra at the site of injection.

In our case study, in addition to these sub-tasks, there were also emergent interruptions competing for the attention of the operator. As a general principle of multi-task performance analysis, when subjects are faced with an interruption, they may perform poorly if the interruption provides no useful information, consuming cognitive resources needed to perform the task. Such irrelevant distractions impose a type of cognitive burden known as extraneous load, which is associated with unnecessary information and should be reduced or eliminated (refer to section 1.2.2 for details). However, in those cases where extraneous load cannot be eliminated or reduced, such

[6]For the sake of consistency, the c-arm was placed in the same position, facing the mannequin throughout the experiments (Figure 5.4). Subjects, however, could switch between the lateral and AP views using a dedicated footpedal.

as of adverse events in the OR, operators must illustrate necessary mental skills to process the extraneous load with minimum cognitive work. Failure to do so may lead to cognitive overload, which in turn, severely disturbs the efficiency and effectiveness of information transfer [40]. In this study, this phenomenon was shown to be true as introducing interruptions significantly increased the task completion time. Although not significant, surgeons tended to administer a larger dose of radiation via fluoroscopy and perform less accurately when faced with interruptions. This also suggests that training for technical skills does not necessary contribute to a better performance when encountered with interruptions in the OR. This was evident in our results as trainees tend to perform better after a short session of training, yet their performance significantly degraded in the presence of interruptions. In addition to the objective results discussed above, our subjective analysis of mental workload also indicates the importance of fully immersive environments with realistic interruptions to develop necessary cognitive skills. In particular, trainees reported greater mental workload when faced with interruptions. Their report of mental workload also correlated negatively with their performance.

The second notion is maintaining patient-safety. Performing vertebroplasty requires navigating the trocar to the target while applying short bursts of X-rays to receive more accurate visual information about the relative position of the tool and the target. Cognitively speaking, this is a mentally competing task: increasing the number of X-ray images, on one hand, may improve the navigation performance, while on the other, compromises patient safety. In such cases, the junior surgeons must choose some task strategy, despite the trade-off between the desired effect of obtaining X-rays on patient safety and navigation performance. In other words, it is necessary to train residents to learn how to maximize patient safety by applying the least possible amount of radiation to generate the best possible mental image of the anatomy and trocar. Each surgeon may choose a different trade-off: cautious individuals may impose lower ra-

diation but might be less accurate and/or slower on the targeting task. Conversely, some may wish to perform more accurately and/or perform the targeting task more quickly, and one strategy to do so would be to perform better on the perceptual task by increasing the duration of X-ray exposure. Of course, these strategy differences do not need to be dramatic shifts but within a tolerable range. Such ability cannot be gained easily by following the traditional *see one, do one, teach one* paradigm. The ability to practice such a unique skill, i.e. performing an image-guided targeting task without substantially compromising patient safety, is one of the key benefits of the proposed fully-immersive simulation environment. Moreover, while obtaining visual-spatial information is necessary to succeed in the navigation task, it is nevertheless subjective, as some junior surgeons have developed higher skills of perception and spatial reasoning than others during their training. Hence, some may need to obtain more X-ray images to attain the same level of perceptual accuracy or to form a spatial representation of the perceptual-motor task space than others. This is manifested in the form of large standard deviations across subjects as illustrated in Table 5.1.

To summarize our findings:

i Our objective and subjective results suggest that training residents to obtain technical skills, though necessary, do not necessarily prepare them to cope with interruptions and different crises scenarios in the OR. This, however, can be practiced with simulators within fully-immersive environments;

ii Maintaining patient-safety is a cognitively challenging task and need to be included in the current training curriculum. Such skills cannot be obtained through traditional training approaches, which further emphasize the practicality of simulators as a training tool;

5.4 Conclusion

Cognitive load is a multifaceted construct influenced by surgeon's technical and non-technical skills, the surgical task, and the OR's environments and surrounding circumstances. Cross-cutting approaches combining both procedural and mannequin-based simulation paradigms are proposed to improve both technical and non-technical skills, preparing trainees to cope with extraneous loads such as distractions, interruptions, and crises. In this work, a user study was conducted in a complete medical simulation environment with occurrence of adverse events to investigate the impact of interrupting events on surgeons' individual experience. Our subjective and objective results suggest that interruptions degrade the efficiency and effectiveness of surgical performance, jeopardizing patient safety. Therefore, training surgical residents under mixed-mode scenarios seems necessary to reduce surgical errors caused by cognitive overload.

5.5 Limitations and Future Work

Due to our participants' time restriction, we could not validate the entire surgical workflow that involves formulating the entry point and cement injection, all of which could be potentially affected by interruptions and so needed to be investigated, nor we could deploy longitudinal training to better assess the effect of teaching non-technical skills on surgical performance. Although increased mental load was evident based on our results, a larger sample size is required to better investigate the role of interruptions on surgical accuracy and patient safety.

Future studies involve investigating the entire workflow, conducting longitudinal training for both technical and non-technical skills - individually and combined, and studying the construct validity of the simulator.

Bibliography

[1] Karen Walker-Bone, Elaine Dennison, and Cyrus Cooper. Epidemiology of osteoporosis. *Rheumatic Disease Clinics of North America*, 27(1):1–18, 2001.

[2] Alexandra Papaioannou, Suzanne Morin, Angela M Cheung, Stephanie Atkinson, Jacques P Brown, Sidney Feldman, David A Hanley, Anthony Hodsman, Sophie A Jamal, Stephanie M Kaiser, et al. 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in canada: summary. *Canadian Medical Association Journal*, 182(17):1864–1873, 2010.

[3] Anne C Looker, National Center for Health Statistics (US), et al. *Osteoporosis Or Low Bone Mass at the Femur Neck Or Lumbar Spine in Older Adults, United States, 2005-2008*. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, 2012.

[4] KE Ensrud, DE Thompson, JA Cauley, MC Nevitt, DM Kado, MC Hochberg, AC Santora 2nd, and DM Black. Prevalent vertebral deformities predict mortality and hospitalization in older women with low bone mass. fracture intervention trial research group. *Journal of the American Geriatrics Society*, 48(3):241–249, 2000.

[5] Deborah M Kado, Warren S Browner, Lisa Palermo, Michael C Nevitt, Harry K Genant, and Steven R Cummings. Vertebral fractures and mortality in older

women: a prospective study. *Archives of internal medicine*, 159(11):1215–1220, 1999.

[6] Antonio Krueger, Christopher Bliemel, Ralph Zettl, and Steffen Ruchholtz. Management of pulmonary cement embolism after percutaneous vertebroplasty and kyphoplasty: a systematic review of the literature. *European Spine Journal*, 18(9):1257–1265, 2009.

[7] Wendy T Ploeg, Albert G Veldhuizen, Maurits S Sietsma, et al. Percutaneous vertebroplasty as a treatment for osteoporotic vertebral compression fractures: a systematic review. *European Spine Journal*, 15(12):1749–1758, 2006.

[8] Frank M Phillips, F Todd Wetzel, Isadore Lieberman, and Marrion Campbell-Hupp. An in vivo comparison of the potential for extravertebral cement leak after vertebroplasty and kyphoplasty. *Spine*, 27(19):2173–2178, 2002.

[9] Alexandra Montagu, Archie Speirs, James Baldock, James Corbett, and Margot Gosney. A review of vertebroplasty for osteoporotic and malignant vertebral compression fractures. *Age and ageing*, 41(4):450–455, 2012.

[10] W.S. Halsted. *The Training of the surgeon*. 1904.

[11] Daniel Resnick and John Barr. *Vertebroplasty and kyphoplasty*. Thieme, 2005.

[12] Jeffrey H Barsuk, Elaine R Cohen, Timothy Caprio, William C McGaghie, Tanya Simuni, and Diane B Wayne. Simulation-based education with mastery learning improves residents' lumbar puncture skills. *Neurology*, 79(2):132–137, 2012.

[13] Kent R Van Sickle, E Matt Ritter, and C Daniel Smith. The pretrained novice: using simulation-based training to improve learning in the operating room. *Surgical innovation*, 13(3):198–204, 2006.

[14] M Passiment, H Sacks, and G Huang. Medical simulation in medical education: Results of an aamc survey. *Association of American Medical Colleges. Washington DC*, pages 1–48, 2011.

[15] Atul A Gawande, Michael J Zinner, David M Studdert, and Troyen A Brennan. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery*, 133(6):614–621, 2003.

[16] Lawrence W Way, Lygia Stewart, Walter Gantert, Kingsway Liu, Crystine M Lee, Karen Whang, and John G Hunter. Causes and prevention of laparoscopic bile duct injuries: analysis of 252 cases from a human factors and cognitive psychology perspective. *Annals of surgery*, 237(4):460, 2003.

[17] S Yule, R Flin, S Paterson-Brown, and N Maran. Non-technical skills for surgeons in the operating room: a review of the literature. *Surgery*, 139(2):140–149, 2006.

[18] Selby H Evans. A brief statement of schema theory. *Psychonomic Science*, 8(2):87–88, 1967.

[19] Mark R Wilson, Jamie M Poolton, Neha Malhotra, Karen Ngo, Elizabeth Bright, and Rich SW Masters. Development and validation of a surgical workload measure: the surgery task load index (surg-tlx). *World journal of surgery*, 35(9):1961–1969, 2011.

[20] Rajesh Aggarwal, Oliver T Mytton, Milliard Derbrew, David Hananel, Mark Heydenburg, Barry Issenberg, Catherine MacAulay, Mary Elizabeth Mancini, Takeshi Morimoto, Nathaniel Soper, et al. Training and simulation for patient safety. *Quality and Safety in Health Care*, 19(Suppl 2):i34–i43, 2010.

[21] Peter Safar, Torrey C Brown, Warren J Holtey, and Robert J Wilder. Ventilation and circulation with closed-chest cardiac massage in man. *Jama*, 176(7):574–576, 1961.

[22] Michael S Gordon. Cardiology patient simulator: development of an animated manikin to teach cardiovascular disease. *The American journal of cardiology*, 34(3):350–355, 1974.

[23] David M Gaba and Abe DeAnda. A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. *Anesthesiology*, 69(3):387–394, 1988.

[24] J Hugh Devitt, Matt M Kurrek, Marsha M Cohen, Kevin Fish, Pamela Fish, Alva G Noel, and John-Paul Szalai. Testing internal consistency and construct validity during evaluation of performance in a patient simulator. *Anesthesia & Analgesia*, 86(6):1160–1164, 1998.

[25] Chee-Kong Chui, Jackson SK Ong, Zheng-Yi Lian, Zhenlan Wang, Jeremy Teo, Jing Zhang, Chye-Hwang Yan, Sim-Heng Ong, Shih-Chang Wang, Hee-Kit Wong, et al. Haptics in computer-mediated simulation: training in vertebroplasty surgery. *Simulation & Gaming*, 37(4):438–451, 2006.

[26] Kyle Michael Fargen, Adnan H Siddiqui, Erol Veznedaroglu, Raymond D Turner, Andrew J Ringer, and J Mocco. Simulator based angiography education in neurosurgery: results of a pilot educational program. *Journal of neurointerventional surgery*, 4(6):438–441, 2012.

[27] Axel Thomas Stadie, Ralf Alfons Kockro, Robert Reisch, Andrei Tropine, Stephan Boor, Peter Stoeter, and Axel Perneczky. Virtual reality system for planning minimally invasive neurosurgery. 2008.

[28] Cristian Luciano, Pat Banerjee, Lucian Florea, and Greg Dawe. Design of the immersivetouch: a high-performance haptic augmented virtual reality system. In *Proceedings of the 11th international conference on human-computer interaction, Las Vegas, Nevada*, 2005.

[29] Eva Kassab, Jimmy Kyaw Tun, Sonal Arora, Dominic King, Kamran Ahmed, Danilo Miskovic, Alexandra Cope, Bhamini Vadhwana, Fernando Bello, Nick Sevdalis, et al. blowing up the barriers in surgical training: exploring and validating the concept of distributed simulation. *Annals of surgery*, 254(6):1059–1065, 2011.

[30] Peter Z Fritz, Tim Gray, and Brendan Flanagan. Review of mannequin-based high-fidelity simulation in emergency medicine. *Emergency Medicine Australasia*, 20(1):1–9, 2008.

[31] Steven K Howard, David M Gaba, Kevin J Fish, George Yang, and Frank H Sarnquist. Anesthesia crisis resource management training: teaching anesthesiologists to handle critical incidents. *Aviation, space, and environmental medicine*, 63(9):763–770, 1992.

[32] R Aggarwal, S Undre, K Moorthy, C Vincent, and A Darzi. The simulated operating theatre: comprehensive training for surgical teams. *Quality and Safety in Health Care*, 13(suppl 1):i27–i32, 2004.

[33] Michael Meehan, Dan Morris, Calvin R Maurer Jr, Anuja K Antony, Federico Barbagli, Kenneth Salisbury, and Sabine Girod. Virtual 3d planning and guidance of mandibular distraction osteogenesis. *Computer Aided Surgery*, 11(2):51–62, 2006.

[34] Johanna Pettersson, Karljohan Lundin Palmerius, Hans Knutsson, Ola Wahlstrom, Bo Tillander, and Magnus Borga. Simulation of patient specific cervical hip fracture surgery with a volume haptic interface. *Biomedical Engineering, IEEE Transactions on*, 55(4):1255–1265, 2008.

[35] Ming-Dar Tsai, Ming-Shium Hsieh, and Shyan-Bin Jou. Virtual reality orthopedic surgery simulator. *Computers in biology and medicine*, 31(5):333–351, 2001.

[36] Behzad Eftekhar, Mohammad Ghodsi, Ebrahim Ketabchi, and Saman Rasaee. Surgical simulation software for insertion of pedicle screws. *Neurosurgery*, 50(1):222–224, 2002.

[37] Patrick Wucherer, Philipp Stefan, Simon Weidert, Pascal Fallavollita, and Nassir Navab. Development and procedural evaluation of immersive medical simulation environments. In *Information processing in computer-assisted interventions*, pages 1–10. Springer, 2013.

[38] Patrick Wucherer, Philipp Stefan, Simon Weidert, Pascal Fallavollita, and Nassir Navab. Task and crisis analysis during surgical training. *International journal of computer assisted radiology and surgery*, pages 1–10, 2014.

[39] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183, 1988.

[40] John Sweller, Paul Ayres, and Slava Kalyuga. *Cognitive load theory*, volume 1. Springer, 2011.

Chapter 6

Closing Remarks

In current surgical practice, *image-guidance* has become an essential part of minimally-invasive surgeries, compensating for the lack of direct vision and limited access to the surgical field. In image-guided interventions (IGI), the surgeon makes use of spatiotemporal information to navigate through the patient body and guide the surgical tools and therapeutic devices to the site of operation. This process is composed of smaller subprocesses that involve collecting, tracking, registering, and displaying information. Collecting information usually begins with acquiring 3D tomographic images prior to the surgery. During the surgery, a localizer is used to continuously track the location of the surgical tools and therapeutic devices. The spatiotemporal information provided by the localizer and preoperative images are then registered together, and the result is presented to the surgeon within a coherent coordinate system. Throughout this process, the surgical outcome can be severely degraded or improved depending on the type of information presented to the surgeon (e.g., preoperative vs. interaoperative, visual vs. haptic), the amount of information, and the method of presentation and interaction (e.g., HMDs vs. computer displays, 3D vs. 2D visualizations).

The importance of presentation and interaction in IGI is two-fold: first, in the current state of image-guided systems, information display is the most vulnerable stage

of the process where data can most easily be misinterpreted by the human observer. Second, due to a large cognitive load imposed by the surgery and OR condition, poor modes of visualization and interaction can overload the surgeon's cognitive capacity, leading to surgical errors and/or long operation time. Under such conditions, the surgeon may have to spend too much time and effort to perceive and process information instead of considering the surgery. Good, intuitive modes of visualization and interaction prevent misinterpretation and replace resource-demanding cognitive tasks with simpler perceptual tasks, reducing the use of the surgeon's mental resources. To accomplish this, a strong understanding of the clinical context as well as a thorough knowledge of human visual perception, cognition, and action are required. The issue of image visualization and interaction influences both the technical and non-technical skills of the surgeon, spanning from preoperative planning to surgical training, simulation, and navigation. This thesis is therefore aimed at studying such aspects of human psychology with the goal of improving minimally-invasive interventions. To this end, I investigated the following four neurosurgical applications in the realm of planning, navigation, training, and training for planning.

Chapter 2 involved investigating the role of stereopsis in lowering the morbidity in ETV interventions caused by accidental perforation of the basilar artery. The result of our user studies suggest that, compared with conventional endoscopes, the detection (and therefore avoidance) of the basilar artery on the surface of the third ventricle can be facilitated with the use of stereoendoscopes, increasing the safety of targeting in third ventriculostomy procedures.

Due to their complex anatomy, resecting arteriovenous malformations is an extremely complex procedure, necessitating a careful preoperative planning. However, 2D NC-MRA images - as a standard of care for planning AVM interventions - may not convey necessary information regarding the vessel connectivity or relative depth. Accordingly, in Chapter 3, I introduce a non-photorealistic technique in conjunction with

volume rendering to enhance the contours of vascular structures in NC-MRA images. Our user studies indicate that the proposed method, particularly when combined with stereopsis, may increase the speed and accuracy of understanding the spatial relationship between vascular structures, improving the quality of planning AVM interventions.

Chapter 4 introduced a mixed augmented/virtual reality system by which junior surgeons could practice spatial reasoning and other cognitive skills necessary in planning tumour resection interventions. A number of user studies were conducted to evaluate the proposed AR system against conventional planning environments. The results indicate that AR may improve subjects' performance, particularly novices', in formulating the optimal point of entry and surgical path independent of the sensorimotor tasks performed, demonstrating the role of mixed-reality systems in assisting residents to develop necessary spatial reasoning skills needed for planning brain tumour resection.

In Chapter 5, a complete simulation environment for vertebroplasty procedures including a mannequin and a simulator with haptic and visual feedback was employed to investigate the effect of interruptions in the OR. We suspected that distractions in the OR may overload the cognitive capacity of novice surgeons, hindering their performance. An evaluation study was conducted with 19 junior surgeons in order to provide a qualitative measure of usability while associating between mental workload and surgical performance. Our results indicate that while training surgeons may increase their technical skills, the introduction of crisis scenarios significantly disturbs the performance, emphasizing the need of realistic simulation environments as part of training curriculum.

Overall, the outcome of this work can be summarized as follows:

- Many intra- and preoperative images and imaging techniques are not perceptually optimized for human observers. Providing simple, yet strong visual cues (e.g., stereopsis), or employing computer graphics techniques to accentuate such cues (e.g., cel-shading) can dramatically facilitate information transfer, complement-

ing surgeons' experience and anatomical knowledge.

- Conventional modes of display are not necessarily optimized to convey visual information, particularly for less trained/experienced users. New emerging methods of visualization and interaction such as AR environments can potentially revolutionize the field of minimally-invasive surgery. More research is needed to improve upon these modes of view by overcoming the current hurdles such as poor ergonomics and accommodation-vergence mismatch.

- Despite recent and ongoing advancements in surgical simulation, much research has been devoted to improving surgeons' dexterity, overlooking their non-technical skills. Our studies have shown that residents can benefit immensely from both new visualization modes (e.g., AR), and fully-immersive simulation environments, acquiring necessary cognitive skills such as spatial reasoning and coping with crises.

In conclusion, human factors such as those discussed in this thesis play a key role in image-guided minimally-invasive surgeries. This was illustrated in this thesis by studying different modes of visualization and interaction and their impact on surgical performance. Although being overlooked in many surgical applications, human factors have received wide attention in recent years and has become the subject of research as an important component of minimally-invasive surgeries.

6.1 Future Direction

Future direction may involve expanding the scope of our work to include other interventions and technologies. Studies regarding human visual perception can be extended to include other perceptual cues, benefiting from both visual and tactile sensory channels. For instance, in addition to occlusion and stereopsis, cues such as aerial perspec-

tive and motion parallax can also be incorporated into our visualization pipeline for planning AVM interventions.

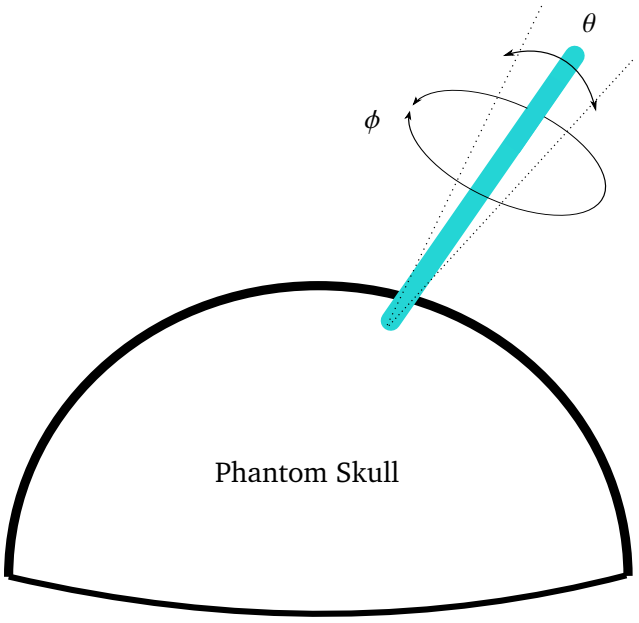
Although we only focused on one specific intervention per project, the underlying technology and evaluation framework can also be employed for (or extended to) similar interventions. For example, our mixed augmented/virtual reality system can be useful for planning AVM interventions. Investigating and possibly adopting these future technologies, such as more advanced augmented-reality solutions or faster GPUs, should also be part of the future work to improve overall patient outcomes.

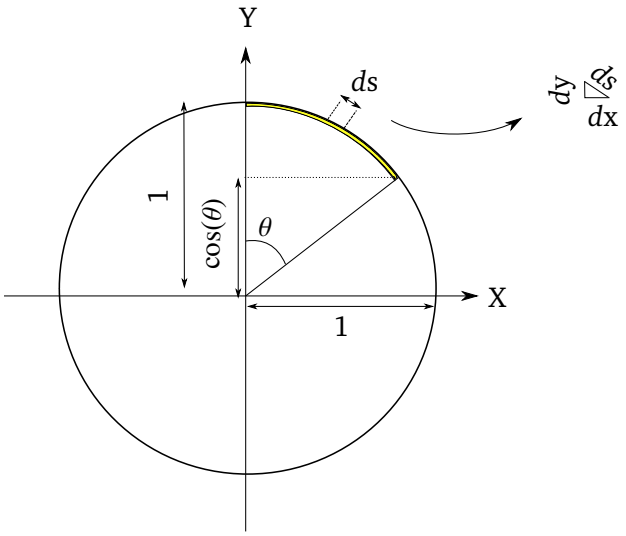
Nevertheless, despite challenges facing studying human factors, such as limited access to clinical participants, it is indeed an exciting, emerging topic of research in the realm of minimally invasive and image-guided interventions, positively influencing the transfer of technology and technological awareness between the laboratory and the operating room.

Appendix A

Derivation of I_p

I_p for LA task based on the rotational error & RT Let the A be space of acceptable x and $Area_A$ be the surface area of A on a unit hemisphere:





$$x^2 + y^2 = 1 \rightarrow y = \sqrt{1 - x^2} \rightarrow dy = \frac{-xdx}{\sqrt{1 - x^2}}$$

$$\begin{aligned}
ds^2 &= dy^2 + dx^2 \\
&= \frac{x^2\, dx^2}{1 - x^2} + dx^2 \\
&= \frac{dx^2}{1 - x^2} \\
&\rightarrow ds = \frac{dx}{\sqrt{1 - x^2}} \\
S &= \int_{\cos(\theta)}^{1} y \,\mathrm{d}s \\
Area_A &= \int_0^{2\pi} \mathrm{d}\phi \int_{\cos(\theta)}^{1} y \,\mathrm{d}s \\
&= \int_0^{2\pi} \mathrm{d}\phi \int_{\cos(\theta)}^{1} \sqrt{1 - x^2} \frac{\mathrm{d}x}{\sqrt{1 - x^2}} \\
&= \int_{\cos(\theta)}^{1} \int_0^{2\pi} \mathrm{d}\phi \,\mathrm{d}x \\
&= \int_{\cos(\theta)}^{1} 2\pi \,\mathrm{d}x \\
&= 2\pi(1 - \cos(\theta)) \\
P(x \in A) &= \frac{Area_A}{\text{surface area of the unit hemisphere}} \\
&= \frac{2\pi(1 - \cos(\mu))}{2\pi} \\
&= 1 - \cos(\mu),
\end{aligned}$$

which implies

$$
\begin{aligned}
ID &= H[P(x)] - H[P(x|x \in A)] \\
&= -\int P(x)\log(P(x))\,\mathrm{d}x + \int_A \frac{P(x)}{P(x \in A)}\log(\frac{P(x)}{P(x \in A)})\,\mathrm{d}x \\
&= -\int P(x)\log(P(x))\,\mathrm{d}x + \\
&\frac{1}{P(x \in A)}\int_A P(x)\log(P(x))\,\mathrm{d}x - \frac{1}{P(x \in A)}\int_A P(x)\log(P(x \in A))\,\mathrm{d}x \\
&= -\int_0^{\pi/2}\int_0^{2\pi}\frac{1}{2\pi}\log(\frac{1}{2\pi})\sin(\theta)\,\mathrm{d}\phi\,\mathrm{d}\theta + \\
&\frac{1}{1-\cos(\mu)}\int_0^{\mu}\int_0^{2\pi}\frac{1}{2\pi}\log(\frac{1}{2\pi})\sin(\theta)\,\mathrm{d}\phi\,\mathrm{d}\theta - \frac{\log(1-\cos(\mu))}{1-\cos(\mu)}\int_0^{\mu}\int_0^{2\pi}\frac{1}{2\pi}\sin(\theta)\,\mathrm{d}\phi\,\mathrm{d}\theta \\
&= \log(2\pi) - \log(2\pi) - \log(1-\cos(\mu)) \\
&= \log(\frac{1}{1-\cos(\mu)}) \\
I_p &= \frac{1}{RT}ID \\
&= \frac{1}{RT}\log(\frac{1}{1-\cos(\mu)})
\end{aligned}
$$

I_p for SD task using the translational error & RT Let $|X|$ = Total number of points uniformly sampled over the skull, $|A|$ = Number of points sampled within the region with a radius of average translational error,

$$
\begin{aligned}
ID &= H[P(x)] - H[P(x|x \in A)] \\
&= -\int P(x)\log(P(x))\,\mathrm{d}x + \int_A \frac{P(x)}{P(x \in A)}\log(\frac{P(x)}{P(x \in A)})\,\mathrm{d}x \\
&\approx -\sum\nolimits_{x\in X}\frac{1}{|X|}\log(\frac{1}{|X|}) + \sum\nolimits_{x\in A}\frac{1}{|A|}\log(\frac{1}{|A|}) \\
&= \log(\frac{|X|}{|A|}) \\
&= \log(\frac{|A| + |\neg A|}{|A|}) \\
I_p &= \frac{1}{RT}ID \\
&= \frac{1}{RT}\log(1 + \frac{|\neg A|}{|A|})
\end{aligned}
$$

Appendix B

Subjective Analysis for Vertebroplasty Experiment

Questions regarding face validity and the training value of simulation environment were answered, respectively, on a scale of 1 (not realistic) to 5 (very realistic) and 1 (strongly disagreed) to 5 (agreed). 19 novice surgeons responded to these questions with the following means and standard deviations:

Category	Statement	$\mu \pm \delta$
Training phase	How realistic was the visualization of fluoroscopy?	4.41±0.32
	How realistic was the haptic feedback (e.g., bone vs. tissue)?	3.82±0.76
	Was the visual representation of the surgical trajectory helpful to get familiar with the simulation environment?	4.78±0.19
	Was the visual representation of the surgical trajectory helpful to get familiar with the surgical workflow?	4.62±0.35
Crises simulation phases	How realistic was the simulation environment?	4.33±0.44
	Would you act the same in real clinical case?	3.83±1.21
	How realistic were the medical instruments?	4.47±0.48
	How realistic was the movement of the instrument?	3.85±0.86
	How realistic was the function of the instrument?	4.28±0.63
	How realistic was the fluoroscopy simulation with footpedal?	4.58±0.43
Crisis #1	How realistic was the phone call?	4.38±0.55
Crisis #2	How realistic was the patient discomfort?	4.01±0.25
	How realistic was the communication with the anesthetist?	4.31±0.48
Training value of simulation	The simulation environment is suitable for OR team training.	4.82±0.25
	The medical training environment represents a real OR.	4.46±0.69
	The simulation environment is suitable for education and training.	4.77±0.51

Curriculum Vitae

Name:	Kamyar Abhari
Post-Secondary Education and Degrees:	University of Manitoba Winnipeg, MB 2005 - 2007 M.Sc. Western University London, ON 2010 - 2014 Ph.D.
Honours and Awards:	OGS 2012-2013 NSERC-CAMI 2011 - 2013 WGRS 2010 - 2014 AECAI (Best Presentation) 2012
Related Work Experience:	Lead Teaching Assistant, Teaching Assistant Western University, University of Manitoba 2008 - 2012

Publications:

Journal Publications:

Wucherer, P.*, Stefan, P.*, **Abhari, K.***, Fallavolitta, P., Weigl, M., Weidert, S. , de Rib-

aupierre, S., Eagleson, R., Peters, T., Navab, N., Vertebroplasty Performance on Simulator for 19 Surgeons Using Hierarchical Task Analysis, Submitted to *IEEE Transaction on Medical Imaging*, (2014) - ***equal contribution**

Abhari, K., , Baxter, J., Chen, E., Khan, A., Wedlake, C., Peters, T., de Ribaupierre, S., Eagleson, R. "Training for Planning Tumour Resection: Augmented Reality and Human Factors", Revised and Resubmitted, Submission No. TBME-00725-2014, *IEEE Transactions on Biomedical Engineering*, (2014)

Abhari, K., Baxter, J., Khan, A., Peters, T., de Ribaupierre, S., Eagleson, R. "Visual Enhancement of MR Angiography Images to Facilitate Planning of Arteriovenous Malformation Interventions", Revised and Resubmitted, Submission No. TAP-2014-0021.R1, *ACM Transactions on Applied Perception*, (2014)

Abhari, K., de Ribaupierre S., Peters T., Eagleson R. "Evaluation of a VR and Stereo-Endoscopic Tool to Facilitate 3rd Ventriculostomy", Studies in Health Technology and Informatics 163:1-7, (2011)

Lecture Notes:

Abhari, K., Baxter, J., Chen, E., Khan, A., Wedlake, C., Peters, T., de Ribaupierre, S., Eagleson, R. "Use of a Mixed-Reality System to Improve the Planning of Brain Tumour Resections: Preliminary Results" Lecture Notes in Computer Science, Augmented Environments for Computer-Assisted Interventions, Springer Berlin Heidelberg, 7815: 55-66, (2013)

Abhari, K., Baxter, J., Chen, E., Khan, A., Wedlake, C., Peters, T., de Ribaupierre, S., Eagleson, R. "The Role of Augmented Reality in Training the Planning of Brain Tu-

mor Resection" Lecture Notes in Computer Science, Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions, Springer Berlin Heidelberg, 8090: 241-248, (2013)

Conference Podium Presentations:

Abhari, K., Baxter, J., Chen, E., Khan, A., Wedlake, C., Peters, T., de Ribaupierre, S., Eagleson, R. Development and Evaluation of an Augmented-Reality Training System for Planning Brain Tumour Resection Interventions, Proceedings of Medical Image Understanding and Analysis, (2013)

Abhari, K., Baxter, J., Eagleson, R., T., de Ribaupierre, S. "Perceptual enhancement of arteriovenous malformation in MRI angiography displays" In SPIE Medical Imaging, pp. 831809-831809. International Society for Optics and Photonics, (2012)

Abhari, K., de Ribaupierre S., Peters T., Eagleson R. "Does stereo-endoscopy improve neurosurgical targeting in 3rd ventriculostomy?" In SPIE Medical Imaging, pp. 796615-796615. International Society for Optics and Photonics, (2011)

Abhari, K., de Ribaupierre S., Peters T., Eagleson R. "Evaluation of a VR and Stereo-Endoscopic Tool to Facilitate 3rd Ventriculostomy", Medicine Meets Virtual Reality, (2011)

Second Author:

Rajch, M., **Abhari, K.**, Stirrat, J., Ukwatta, E., Cantor-Rivera, D., Li, F., Peters, T., and White, J. Distribution of guidance models for cardiac resynchronization therapy in the setting of multi-center clinical trials In: SPIE Medical Imaging. International Society for Optics and Photonics, (2014)

Baxter, J., **Abhari, K.**, Chen, E., Peters, T. Multidimensional Transfer Function Definition via Self-Organizing Mixture Networks, In Progress, to be submitted to IEEE Transactions on Visualization and Computer Graphics (TVCG), (2014)